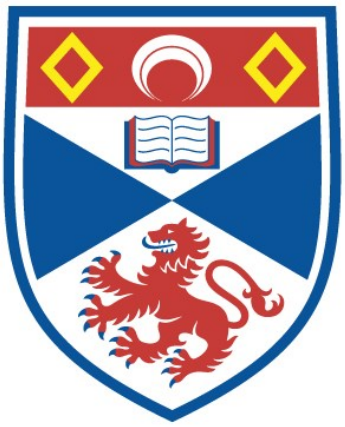
THE EXPRESSION AND PURIFICATION OF RUKH61 AND MB

Cath Stolton

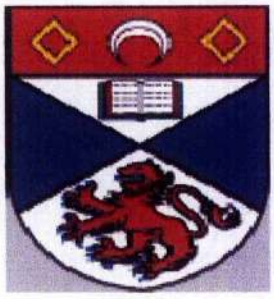
A Thesis Submitted for the Degree of MPhil
at the
University of St Andrews



2000

THE EXPRESSION AND PURIFICATION

OF Rukh61 AND Mb



UNIVERSITY OF ST ANDREWS

DEPARTMENT OF CHEMISTRY

CATH STOLTON

THESIS SUBMITTED FOR THE

DEGREE OF MPHIL

1998

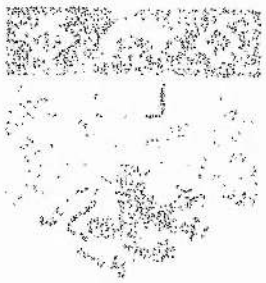


Table of Contents

Abbreviations

Akt/PKB	serine-threonine kinase
AMP	ampicillin, an antibiotic
Arg	arginine
Asp	aspartic acid
bFGF	basic fibroplast growth factor
BSA	bovine serum albumin
CAP	chloramphenicol, an antibiotic
	DNA
DTT	dithiothreitol
EGF	epidermal growth factor
	electrophoresis
Glu	glutamic acid
Gly	glycine
GST	glutathione S-transferase
His	histidine
IGF-1	insulin-like growth factor
IPTG	isopropyl-1-thiol-β-D-galactoside
IRS-1	insulin receptor substrate
Leu	leucine
LMW	low molecular weight
NaAc	sodium acetate
NaCl	sodium chloride
NaOH	sodium hydroxide

NMR	nuclear magnetic resonance
P1 solution	50 mM Tris-HCl, pH 8.0, 10 mM EDTA, 100 µg/ml RnaseA, 4°C
P2 solution	200 mM NaOH, 1% SDS
P3 solution	3.0 M potassium acetate, pH 5.5, 4°C
PBS	phosphate buffer
PDGF	platelet-derived growth factor
PEG	polyethylene glycol
PEST	proline, serine and threonine rich region in a protein
pGEX	protein expression system
PH domain	pleckstrin homology domain
Phe	phenylalanine
PI3-K	phosphoinositide 3-kinase
PLP	phospholipid phosphatase
PMSF	phenylmethanesulphonyl fluoride
Pro	proline
PtdIns(3, 4)P_2	phosphatidylinositol (3, 4) diphosphate
PtdIns(3, 4, 5)P_3	phosphatidylinositol (3, 4, 5) triphosphate
QBT solution	450 mM NaCl, 50 mM MOPS, pH 7.0, 15% isopropanol, 0.15% Triton X-100
QC solution	1.0 M NaCl, 50 mM MOPS, pH 7.0, 15% isopropanol
QF solution	1.25 M NaCl, 50 mM Tris-HCl, pH 8.5, 15% isopropanol
SDS	sodium dodecyl sulphate

SDS-PAGE	sodium dodecyl sulphate – polyacrylamide gel
Ser	serine
SH3/SH2	src homology 3 domain/ src homology 2 domain
SOC	complex buffered medium for growth of competent
STE	sodium chloride, Tris, and EDTA buffer solution
TE	10 mM Tris-HCl, pH 8.0, 1 mM EDTA
Tris	Tris(hydroxymethyl)aminomethane
Val	valine

Acknowledgements

Many thanks to everyone in the Homans and Naismith groups, especially Jim, Louise and Achim, for their help and support over the last year. Thanks as well to Vladimir and Julia for the proteins which I'm working on, and for all of the information and help they have given.

Abstract

The protein Rukh61 was expressed in the fusion vector pGEX-4T-1 and the recombinant organism grown in Luria broth and minimal media. Ruk is expressed early in rat development, in a variety of tissues. Ruk is of interest because of its interaction with proteins in the signal transduction pathway. It is known to inhibit a PI3-kinase pathway which involves Akt/PKB. The mode of action of this protein is of interest because it may lead to new anti-cancer drugs. The protein was purified using affinity chromatography. Thrombin protease was used to cleave the GST tag from the target protein and affinity chromatography used to remove the GST tag from the solution containing the target protein. Concentration of the protein has been achieved through the use of column chromatography, the freeze drying of samples, and by ammonium sulphate precipitation. Through the course of the work, the protocol for the purification of GST-Rukh61 has been optimised, as has the protocol for the cleavage of the GST tag and the concentration of the protein.

The protein GST-Mb has also been expressed in small quantities. Mb is of interest because it is found in mice in neuronal tissue only and it is only found during the later stages of development. Mb responds to nerve growth factors thus its mode of action, investigated through structural studies, is of interest. It has been transformed into a different *Escherichia coli* (*E. coli*) strain, from that in which it was obtained, to facilitate expression. The Mb can also be cleaved from the GST using thrombin protease. Work on the optimisation of the protocols for Mb is ongoing.

1. Introduction

The overall objective of this project is to optimise the expression and the purification of two unrelated proteins. This would then lead to the preparation of samples for NMR analysis and structural studies.

The proteins of interest are Rukh61 and Mb. Rukh61 is a shortened version of Ruk, it is more amenable to NMR studies through being smaller whilst being able to give an indication of the 3D structure of Ruk. Structure determination by NMR is currently limited to proteins of <300 (~30kDa) amino acids[44] because the number of resonances in larger proteins gives an overcrowded, overlapped spectrum which is very difficult to assign. Ruk is able to inactivate phosphatidylinositol-3-kinase (PI3-kinase) signal transduction pathways,[22] which in turn leads to the inactivation of Akt/PKB and the mechanisms that are dependent on it, mechanisms such as the activation of glycogen synthesis, and the regulation of cell survival.[29]

Mb is found in neuronal tissue of chicks, humans, mice and rats, and is only found in the later stages of development.[22] The amino acid sequence shows a repeat unit with a region in-between the repeating units which sequence analysis indicates may be a loop structure. Further sequence analysis of this loop region has led to the conclusion that it may be more than one small loop.[15]

The pGEX GST[1,2] expression system is used to obtain high level expression and rapid purification of GST tagged proteins. Genes encoding Mb and Rukh61 have been cloned into the vector pGEX-4T-1 to produce the plasmid with antibiotic resistance so

that only the desired protein grows. Ampicillin (AMP) resistance is conferred by the pGEX vector and chloramphenicol (CAP) resistance is conferred by the *E. coli.* strain BL21(DE3)pLysS.

A vector map of the pGEX-4T-1 vector is shown in Figure 1.

The site labelled thrombin is the recognition site and cleavage is seen to occur between Arg and Gly.

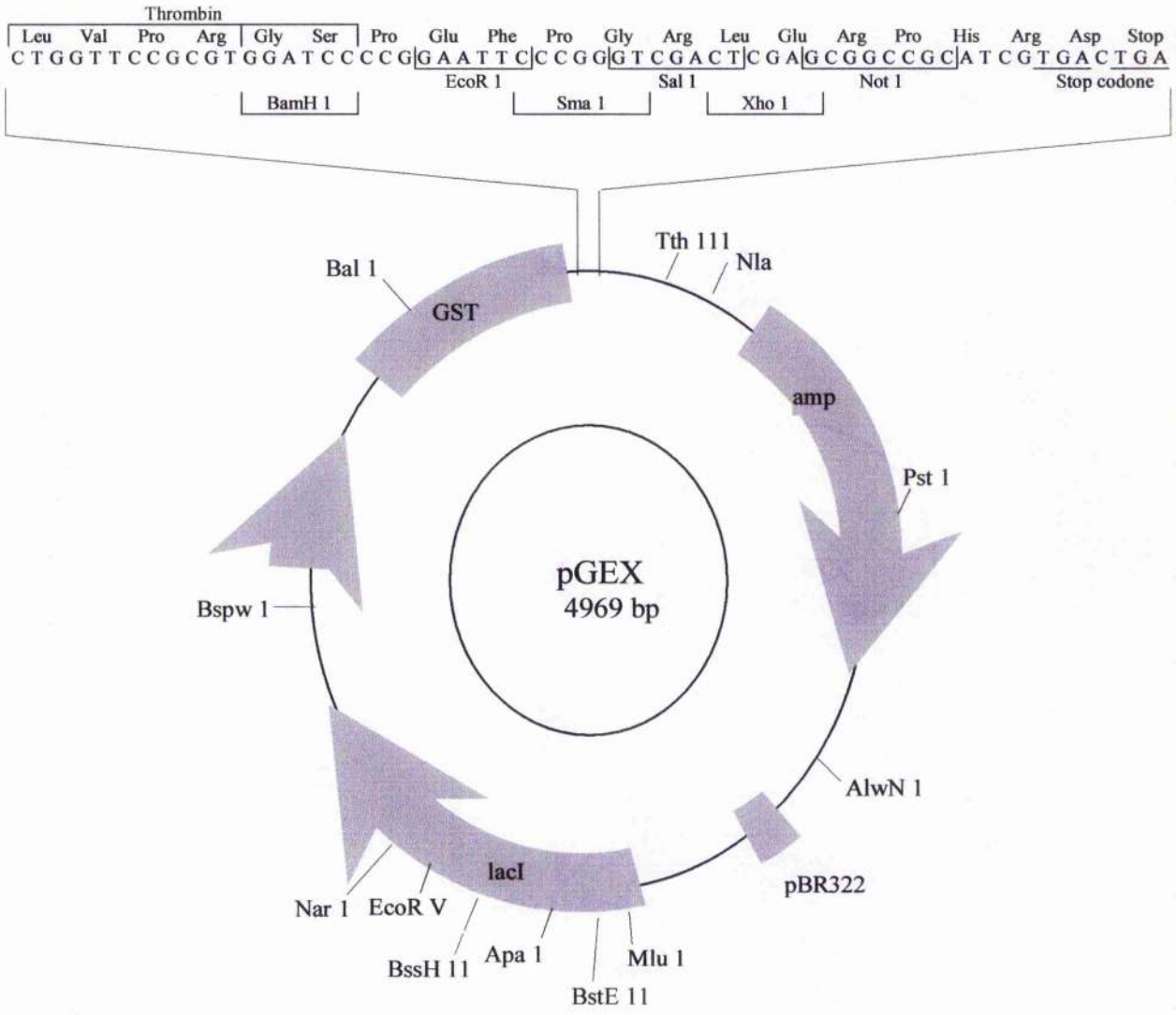


Figure 1: Vector map of pGEX-4T-1

This diagram shows the genetic map for the pGEX vector with thee sites of interest, i.e. the thrombin site and the positions of restriction sites that are unique within this vector.

1.2 Introduction to Ruk

The protein Rukh61 is a shortened version of the protein Ruk and is the heart variant. Ruk is involved in the phosphoinositide 3-kinase (PI3-K) and serine-threonine kinase (Akt/PKB) signal transduction pathway. It deactivates Akt/PKB by inhibiting PI3-K. Akt/PKB was discovered by two groups independently who named it differently,[3,4] to avoid confusion the term Akt/PKB shall be employed throughout this thesis.

Ruk Structure

There are three principle forms of Ruk mRNA, resulting in three variations of the protein: long, medium and short.[22] Figure 2 shows the schematic structure and possible splicing forms of Ruk. The protein under investigation in this work is the shortened form of Ruk, known as Rukh61. This is the short protein shown in Figure 2 with a coiled-coil domain only.

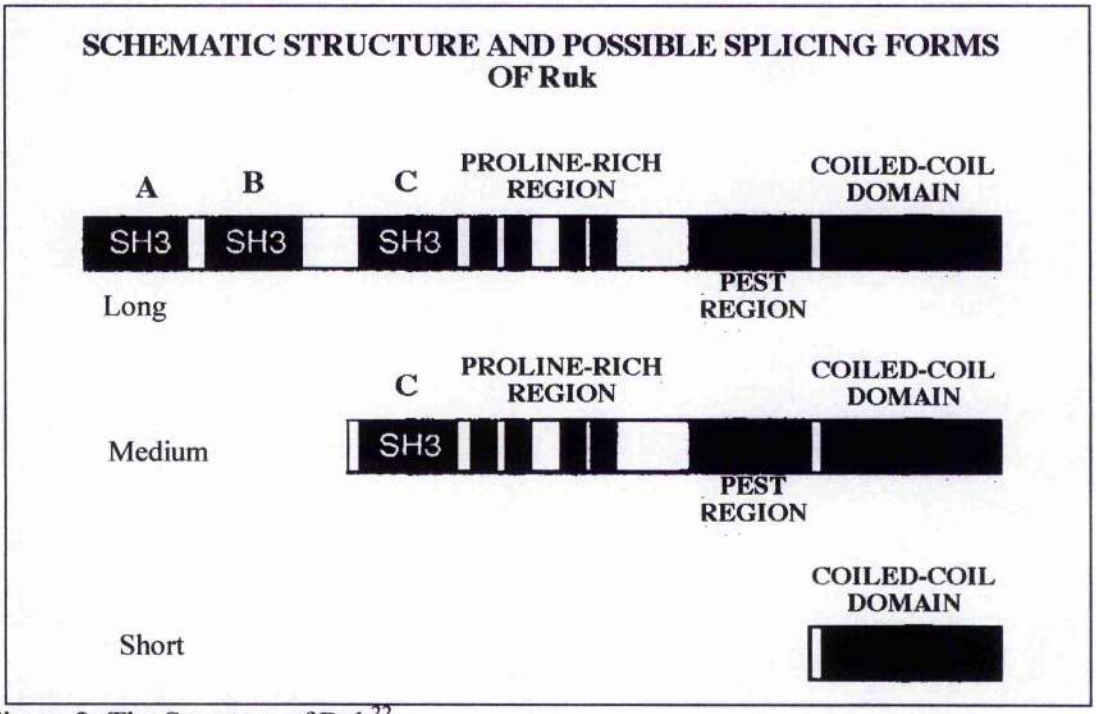


Figure 2: The Structure of Ruk[22]

There are three principle forms of Ruk proteins. These are shown above and are labelled as long, medium and short, referring to their relative lengths. Rukh61 is the short protein observed in the diagram. This diagram shows the regions that are found in the differently spliced forms of the protein. SH3 domains are found in many signalling molecules and bind selectively to proline-rich sequences. PEST regions are regions rich in proline, serine and threonine.

The carboxy terminal domain of the family of Ruk proteins is conserved throughout the heart, lung and skin variants of the rat protein.[22] This is in contrast to the N-terminal domain containing SH3 domains and a proline rich domain which vary amongst the tissues. The PEST region is a proline, serine and threonine rich region which is usually present in proteins with a high metabolic turnover.[22] SH3 domains contain 55-70 residues and are present in a large number of signalling molecules.[5] The structures of SH3 domains that have been solved are similar to each other.[5] They are known to selectively bind to proline-rich sequences, such as that present in Rukh61. SH3 domains mediate protein-protein interactions akin to those required in some signal transduction pathways.[5] Proteins can have multiple SH3 domains, for example, the long form of Ruk as shown in Figure 2, allowing the potential to form a cluster of

several distinct ligands. Serine or threonine phosphorylation adjacent to the proline rich ligand can influence SH3 domain interactions.[6] The prolines have two roles in the interaction between an SH3 domain in an adapter protein, such as GRB2, and a proline-rich domain in another. The proline-rich sequence assumes an extended conformation that allows extensive contacts with the SH3 domain.[22] A subset of these proline sequences is thought to fit into binding pockets on the surface of the SH3 domain.[7] SH2 domains bind specific phosphotyrosyl residues on activated receptors.[7] Serine phosphorylation of an SH3 recognition site results in the uncoupling of signalling proteins.[7]

C-terminal domain of Ruk

It has been proposed, on the basis of structure prediction, that the C-terminal domain of Ruk has a coiled-coil structure.[22] This type of structure is formed from two amphipathic α-helices wrapped around each other. An amphipathic α-helix is one in which the hydrophilic side chains are on one side of the helix and the hydrophobic side chains are on the opposite. Within a coiled-coil structure the helices wrap around each other with the hydrophobic residues projecting from one helix interdigitating into the gaps between the hydrophobic side chains of the other helix along the contact surface. The amino acid sequence of each amphipathic helix contains a characteristic 7-residue "heptad" repeat, in which the hydrophobic residues are predominately at positions a and d. A study of coiled-coils has shown that it takes at least three heptads or six helical turns to cause a molecule to adopt a coiled-coil conformation in aqueous media.[8]

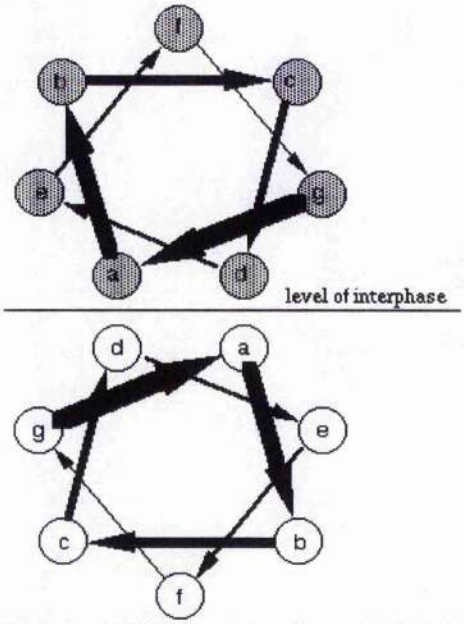


Figure 3: Illustration of a coiled-coil.
A coiled-coil occurs when two amphipathic helices wrap around each other and the hydophobic residues at positions a and d in the diagram interdigitate into the gaps in the other helix along the contact surface.

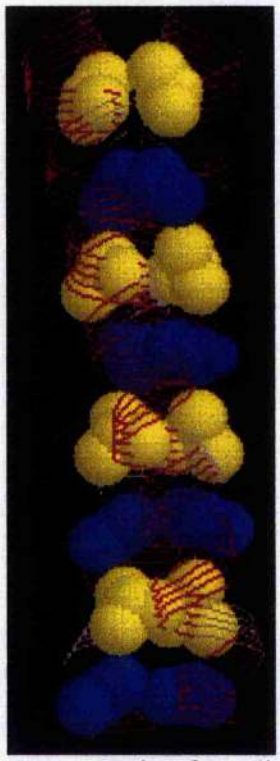


Figure 4: An example of a coiled-coil.
GCN4 leucine zipper.[11]

Figure 3 shows an illustration of the amphipathic nature of helices.[11] The coiled-coil motif is quite tolerant to sequence variation as is shown by the variety that appear in nature.[9] The coiled-coil structure is also found in the Gal4 dimer and the leucine-zipper dimer family. Leucine zippers are a widely observed structural motif which promotes homo- and hetero-dimerisation in some DNA-binding proteins.[8] Dimerization occurs through the formation of a short parallel coiled-coil of α-helices. The zipper is characterised by a heptad repeat of leucine residues at position d. Another heptad repeat of hydrophobic residues occurs at position a. These residue pack against each other in the coiled-coil structure.[10] In many natural and artificial leucine zippers asparagine, in position a, enforces a parallel orientation of strands. An example of a leucine zipper is GCN4. This is shown in Figure 4 above.[11]

Other examples from nature include the coiled-coil domain in kinesin which is

structurally important for co-ordinating the activities of kinesin heads during processive movement.[12] Tropomyosins are a family of actin binding proteins found in eucaryotic cells. Tropomyosin is a coiled-coil protein along its length apart from the ends, it binds to actin in such a way as to align the ends.[13] Multimerin is a 1228 amino acid, soluble protein that is found in platelets and endothelium blood vessels, the central portion of its sequence shows a coiled-coil structure. The role of this domain has not yet been determined, but it has been assumed to be a site for interchain associations.[14]

In the case of the C-terminal of Ruk the proposed coiled-coil structure is underlined in the sequence of Rukh61 shown. This was obtained from Buchman[22] and shows the coiled-coil to be between residues 53-110. A consensus of eight types of structure prediction was carried out[15] using Network Protein Sequence Analysis at IBCP. This gave a coil structure for residues 1-11, 15-34 and 40-41; a sheet structure for residues 12-14 and a helical structure for residues 35-39 and 42-110. This helical structure could make up the coiled-coil. Analysis using the method developed by Lupas *et al.*[16] for the prediction of coiled-coils from protein sequences gave a probability of > 0.99 of a coiled-coil between residues 53 and 105, which is again close to the residues shown. The use of PHD (Profile fed neural network systems from HeiDlberg) gave rise to similar results with a prediction of helices for residues 51-107.[17]

Coiled-coil sequences in nature have many variations with a leucine molecule at position d being, in general, retained. The coiled-coil domain of Rukh61 shows approximately 40% similarity with a range of leucine zipper domains, including the coiled region of GCN4, with the leucine zipper investigated by Jelesarov[18] and with a

natural bZIP protein investigated by Moitra *et al.*[19]

Rukh61 contains 110 amino acids, the sequence being:

MAAASSGPAS LSSVASSPMS SSLGTAGQRA SSPSLFSAEG KAKTESAVSS QAAIEELKMQ VRELRTIIET MKDQQKREIK QLLSELDEEK KIRLRLQMEV NDIKKALQSK

Rukh61 has a molecular weight of 11.896 kDa, to this is added a GST tag of 26 kDa to enable the easy purification of the protein, giving a combined molecular weight, before cleavage, of 38 kDa. Rukh61 has been cloned into the vector pGEX-4T1 by Buchman[22] and was expressed in *Escherichia coli* (*E. coli*) strain BL21(DE3)pLysS.

The computer programme 'Protean'[20] was used to determine the statistics of Rukh61 from its amino acid sequence. Table 1 shows results of the analysis carried out.[15]

Molecular Weight	11896.00 Da
I microgram =	84.062 pMoles
Molar Extinction coefficient	0
1 A(280) =	*
Isoelectric point	9.355
Charge at pH 7	2.928

Table 1: The analytical data of Rukh61 predicted from the amino acid sequence.
*Rukh61 has no amino acids which show at an absorbance of 280nm.

PI3-K

PI3-K is a heterodimer made up of 110 kDa and 85 kDa subunits.[21,22] These subunits are referred to as p110 and p85, the p110 subunit containing a catalytic domain which confers lipid kinase activity and a binding site for wortmannin, an inhibitor of PI3-K.[21] The p85 subunit contains two proline rich motifs which are responsible for binding to src homology 3 (SH3) domains,[35] these are protein-protein interaction domains. It also contains two src homology 2 (SH2) domains.[21] The p85 subunit shows no kinase activity.[21] The p85 subunit can be further subdivided into a p85α and a p85β domain.

Role of Ruk and Interaction with PI3-K

It has been shown by Buchman[22] that Ruk RNA is expressed in the first fifteen stages of new-born rat development, with the level of expression dropping as development continues. It has also been shown that in different tissues, for example skin, muscle and heart, different versions of Ruk RNA are expressed. Buchman has also obtained evidence that Ruk can, *in vitro*, interact with proteins involved in the signal transduction pathway.[22] This was found through an investigation of the SH3 domains of Ruk and their binding to proteins with SH3 and SH2 domains, especially the N-terminal SH3 domain of GRB2. GRB2 contains an SH2 domain which binds to phosphotyrosine residues and two SH3 domains.[22] Experiments *in vitro* have shown that the SH3 domains of Ruk bind well to p85α which is a regulatory subunit of PI3-K.[21,22] This complex between the subunit and Ruk also exists *in vivo*, as has been shown through immunoprecipitation experiments by Buchman.[22] This also showed that the p85α subunit and not the p85β subunit complexes with Ruk. Further experiments showed that the PI3-K activity is inhibited by the presence of Ruk.[22]

Continued experiments by Buchman using sensory neurons (these are dependent on neurone growth factor which is inhibited by Ruk) with the plasmid to produce Ruk, have shown that it only needs the SH3 and/or the proline rich domain of the protein to inhibit survival of the neurons. This means that the coiled-coil domain is not used to form the complex with p85α. Experiments on these neurons also showed[22] that the effect of Ruk is reversible by activating downstream members of the pathway, that is, the injection of active Akt/PKB or PI3-K will allow the survival of neurons.

Akt/PKB

Akt/PKB is a serine-threonine kinase, which contains a pleckstrin homology (PH) domain at the N-terminal, residues 1-106.[23] This domain regulates the activation of Akt/PKB through the binding of the D3-phosphorylated phosphoinositides that are produced by PI3-K.[24] The phosphorylation of Akt/PKB influences its activation and the PH domain may also influence the activation by promoting dimerization and the formation of Akt/PKB protein complexes.[23] It is known that Akt/PKB activation is part of a PI3-K dependent signalling pathway, because when treated with two unrelated PI3-K inhibitors, such as wortmannin and LY294002, the Akt/PKB is not activated. The data collected by Burgering and Coffer,[24] using an epitope-tagged version of Akt/PKB, led them to the conclusion that PI3-K activation is essential for Akt/PKB activation to occur. Akt/PKB has been found to show increased activity after treatment with platelet-derived growth factor (PDGF), epidermal growth factor (EGF), insulin and basic fibroplast growth factor (bFGF).[24] The intracellular signalling pathways, through which growth factors promote survival of neurons in the central nervous system, are not well characterised. Pathways which include PI3-K are important for the survival of several cell lines,[25] however the modes of operation are

unclear. Two targets of PI3-K are Akt/PKB and p70 S6-kinase ($p70^{S6k}$) which contribute to the insulin-like growth factor 1 (IGF-1) mediated survival of cerebellar neurons.[32] It has been established that IGF-1 and insulin activate PI3-K in cerebellar neurons, this occurs as they induce binding of PI3-K to the receptor-associated protein IRS-1 (insulin receptor substrate 1).[32] IRS-1 mediates the activation of PI3-K and IGF-1 and insulin increase the activity of PI3-K.[32] Insulin and the IGF-1 also activate $p70^{S6k}$, but the phosphorylation of $p70^{S6k}$ is blocked by the PI3-K inhibitor wortmannin, which suggests that $p70^{S6k}$ activation is dependent upon PI3-K.[30,32]

Figure 5 shows the pathway of activation of Akt/PKB. It starts at the insulin receptors and goes through the activation of the p110 subunit of PI3-K by the p85 subunit or by activated RasGTP.[35] RasGTP is activated by SOS and Shc-GRB2.[35] Then the subsequent phosphorylation of Akt/PKB occurs.[31] LY294002 is an inhibitor of PI3-K and is shown as not reacting in the activation pathway. The reaction of phosphatidylinositol (4,5)-diphosphate [PtdIns(4,5)P_2] with PI3-K leads to the production of PtdIns(3,4,5)P_3. The reaction of PtdIns(3,4,5)P_3 with PLP (phospholipid phosphatase), leads to the production of PtdIns(3,4)P_2, as is shown in the diagram. Both of these products bind to the PH domain of Akt/PKB and stimulate its activation through the phosphorylation of Thr308 and Ser473.[28] The activation of Akt/PKB involves a change in conformation.[28,30]

The diagram shows that the activation of Akt/PKB is Ras-dependent,[22,35] suggesting that it may be a member of a signalling pathway. The data obtained by T. F. Franke *et al* [26] suggests that Akt/PKB functions in a pathway that is Ras-dependent but MAP kinase independent, such a pathway involving MEKK and JNK was recently

described[27] so there is a precedent. The other pathway shown in Figure 5 involves Ras interaction with serine/threonine kinases of the Raf family. When Raf has been activated by activated RasGTP it phosphorylates, and thus activates, MEK. This pathway is interrupted by PD90859 as shown.[31]

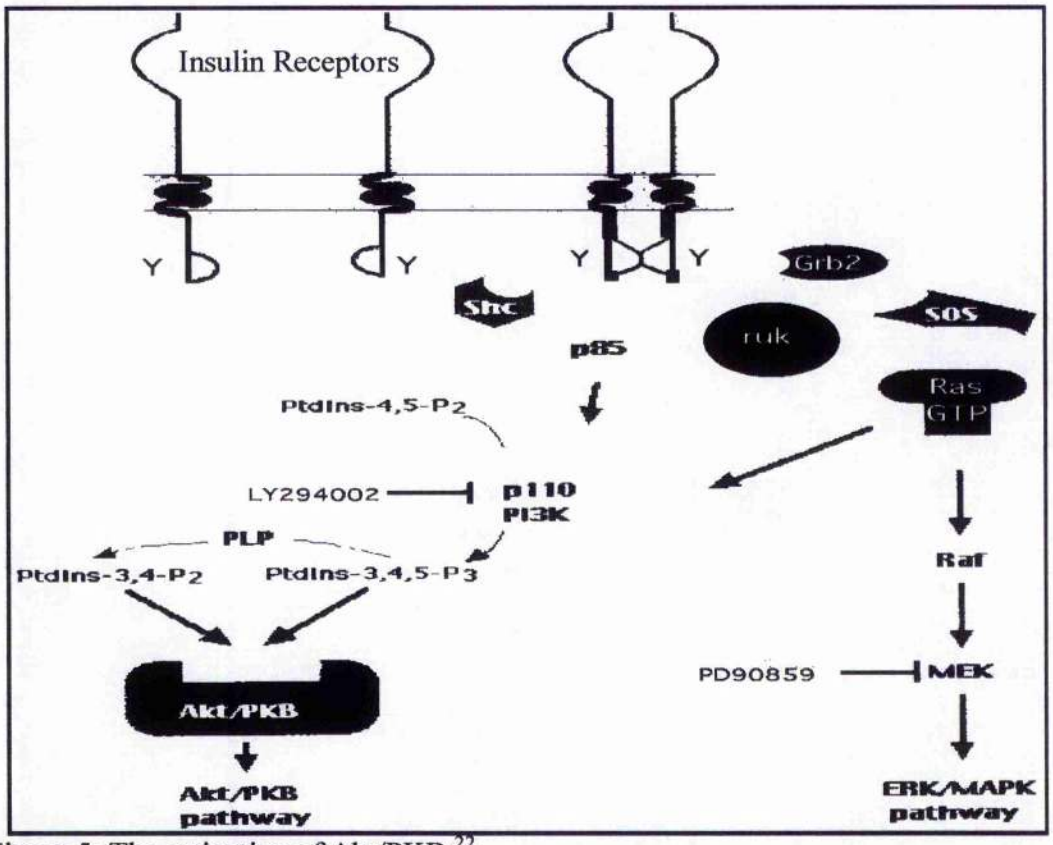


Figure 5: The activation of Akt/PKB.[22]

This shows the path of activation of Akt/PKB from the stimulation by the insulin receptors of the p85 subunit. The p110 subunit of PI3-K is activated by interaction with either the p85 subunit or with activated RasGTP.[35] RasGTP is activated by SOS and Shc-GRB2. ruk can also interact and is capable of disrupting the activation process. LY294002 is an inhibitor of PI3-K and is shown as not reacting in the activation pathway. PtdIns(4,5)P_2 reacts with PI3-K to produce PtdIns(3,4,5)P_3. This interacts with PLP to give PtdIns(3,4)P_2. Both of these bind to the PH domain of Akt/PKB and stimulate activation through the phosphorylation of Thr308 and Ser473. The activation of Akt/PKB involves a change in conformation.[28,30] The other pathway shown involves Ras interaction with serine/threonine kinases of the Raf family. Once Raf has been activated it phosphorylates and activates MEK. This pathway is interrupted by PD90859 as shown.[31]

Burgering and Coffer[24] also undertook a study of the mechanism of Akt/PKB activation to investigate whether Akt/PKB is regulated by phosphorylation, this Akt/PKB was induced by PDGF and inhibited by treatment with wortmannin.

Analysis of the Akt/PKB showed that phosphorylation was mainly on Ser473. The phosphorylation appears to be essential for Akt/PKB kinase activity as treatment of active Akt/PKB with phosphatase removed the ability of Akt/PKB to be phosphorylated. One response to the activation of Akt/PKB is the activation of $p70^{S6k}$. This provides evidence that Akt/PKB is an element of a growth factor induced signal transduction pathway and implies that the PI3-K pathway is sensitive to oncogenic conversion, and thus may be important in the development of human cancer.[24,28] PI3-K and Akt/PKB have been implicated in neuronal differentiation and survival.[25] PI3-K has a role in the mediation of NGF-induced neurogenesis and in the survival of PC12 cells. In these cells PI3-K activity is controlled through the Shc-binding site on Trk, which is not required for NGF-mediated cell survival.[29] It appears that PI3-K is a member of a pathway that collaborates with other pathways in the regulation process. Akt/PKB can be activated by phosphatidylinositol-3,4-biphosphate, which is a secondary messenger generated by PI3-K activity.[29] Inactive Akt/PKB is able to block KCl- and insulin-mediated survival, thus indicating that Akt/PKB is a key component of cell survival of cerebellar neurons.[29]

There are three main stages in the promotion of activation of Akt/PKB by PI3-K derived messengers. First translocation of the kinase to the membrane occurs, followed by the attachment of the kinase to the membrane by way of the PH domain binding to phospholipid and finally phosphorylation occurs. This is shown in Figure 6, which is a model for the activation of Akt/PKB by PI3-K dependent mechanisms.[30]

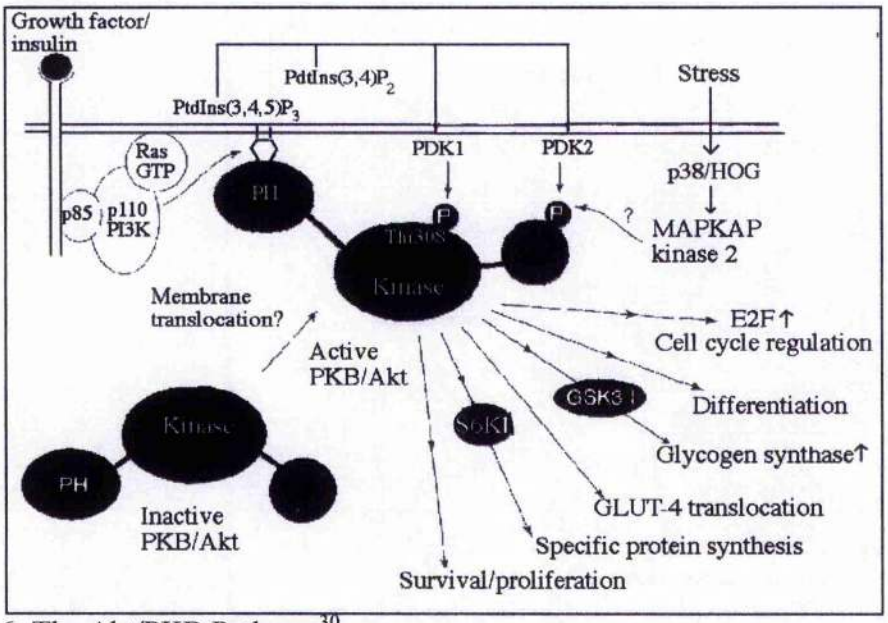


Figure 6: The Akt/PKB Pathway[30]

This diagram shows the pathway of Akt/PKB activation starting from inactive Akt/PKB in the bottom left corner. First translocation of the kinase to the membrane occurs, followed by its attachment to the membrane by way of the PH domain binding to phospholipid [PtdIns(3,4,5)P_3] and then phosphorylation occurs at Thr308 and at Ser473 producing active Akt/PKB. This is shown in more detail in Figure 5. Activated Akt/PKB is involved in many activities as shown in the diagram. Included in these activities is the phosphorylation of glycogen synthase kinase-3, GLUT-4 translocation, cell cycle regulation and, through other pathways, protein synthesis.

Figure 6 shows that following growth-factor receptor simulation,[30] the p110 subunit of PI3-K is activated by interaction with the p85 subunit and activated RasGTP. The activation of PI3-K produces both phosphatidylinositol (3,4,5)-triphosphate [PtdIns(3,4,5)P_3] and PtdIns(3,4)P_2.[28,30] The PH domain of Akt/PKB binds to these phosphoinositides, recruiting Akt/PKB to the plasma membrane and stimulating kinase activity this domain has been shown to be essential for Akt/PKB activation.[31] The full activation of Akt/PKB requires phosphorylation of Thr308 and Ser473, utilising two kinases (PDK1 and PDK2) which are dependent for activation on PtdIns(3,4,5)P_3.[30] The activation of Akt/PKB involves a change in conformation.[28,30]

Akt/PKB is responsible for the phosphorylation of glycogen synthase kinase-3, which not only results in its inactivation, but also in the activation of glycogen synthesis.[32]

This is important as PKBa is amplified and overexpressed in many ovarian neoplasms.[32] Akt/PKB has been found to stimulate glucose uptake and GLUT4 translocation. Further experiments by Dudek *et al*[33] showed that the promotion of survival by insulin requires Akt/PKB. Thus, another function of Akt/PKB is to mediate the effects of IGF-1 on neuronal survival, this may promote the survival of a range of growth factors that activate PI3-K. Activated PI3-K and Akt/PKB have been shown to protect epithelial cells from apoptosis caused by the detachment of adherent cells from their extracellular matrix.[34] Pathways other than PI3-K can lead to Akt/PKB activation. $DiC_{16}PtdIns\text{-}3,4\text{-}P_2$ works even when wortmannin has blocked PI3-K.[34]

Another type of signalling pathways are those which contain receptor tyrosine kinases (RTK's). These signalling pathways have a wide spectrum of functions including regulation of cell proliferation and differentiation, promotion of cell survival and adjustments in cell metabolism.[5] PI3-K is an effector of Ras,[35] Ras proteins are critical components of signalling pathways leading from cell surface receptors to control of cellular proliferation.[35] They are also involved in morphology and differentiation. Ras is able to regulate PI3-K which is important in Ras control of cellular morphology and DNA synthesis.[35]

1.3 Introduction to Mb

Mb is a neurospecific protein which is non-hydrophobic and thus is soluble.[5,22] It has a molecular weight of 9.29kDa[22] which is increased to 35kDa by the addition of a Glutathione S-Transferase (GST) tag. The DNA for the protein has been isolated from chicken neural screening, the protein itself is expressed in a recombinant form. Mb contains 83 amino acids and has been described by Buchman[22] as having two repeating units with a 'novel' loop in-between. The repeat units are underlined in the sequence and the homologies shown. A structure prediction, obtained from Buchman[22] and shown in the Appendix 1, shows predictions for the chick, human, mouse and rat varieties. The one of interest is the chicken form as that is the strain under investigation in this work.

The structure prediction in Appendix 1 is a Momany prediction[22] and shows a probable helix structure for residues 7-18 and 70-80. It also predicts a lop structure at residues 30-35 and 50-53. The probability of a loop structure is conserved in the mouse and rat form, but in the human form whilst a helix is predicted in that region, there is also a high probability of a loop being present.

A consensus of eight types of structure prediction was carried out[15] using Network Protein Sequence Analysis at IBCP. This gave a coil structure in residues 3-10, 17-20, 30-36, 39-46, 50-53, 60-79 and 82-83. Of these residues 30-36 and 50-53 correspond to those seen as loops in the Momany prediction and residues 39-46 are in a region where the Momany prediction is unsure of the structure. The remaining residues are within the repeat units. Another structure prediction using the method developed by

Rost and Sanders[17] also agrees with the Momany prediction, indicating a loop structure for residues 31-34, 40-44 and 49-52.

The amino acid sequence is

MLDTISSQYD SFIYWRMPIP RLELAELEGL GLGQGSLYAP HSKLPEAAQE PGCAEEDSLL PFSSFNFWRA PIASISSFDF DLI

The homology between the N-terminal and the C-terminal regions of Mb can be seen here:

```
N terminal   DTISSQYDSFIYWRMPIPRLELAELEGL
             |::   : ||  :|| ||  :   : : :
C terminal   DSLL-PFSSFNFWRAPIASISSFDFDLI
```

Mb is found in the neuronal tissue of mice, but not in the spleen, lung, heart, liver or kidney. It was found by Buchman[22] to be regulated by the development of the mice. There is a sharp increase in the expression of Mb in the mouse hindbrain, this occurs during the later stages of development. The Mb responds to nerve growth factors in rat cell line PC12 and again shows expression at a late stage of development.

Mb is expressed in the vector pGEX-4T1, which has a thrombin cleavage site and was obtained from Buchman in the *E. coli* strain X90. It has been expressed in this strain and also in the *E. coli* strain BL21(DE3)pLysS.

The computer programme 'Protean'[20] has been used to determine the statistics of Mb from its amino acid sequence. Table 2 shows the results of the analysis carried out.[15]

Molecular Weight	9285.9 Da
I microgram =	107.690 pMoles
Molar Extinction coefficient	15340 ± 5%
1 A(280) =	0.61 mg/ml
Isoelectric point	4.008
Charge at pH 7	- 7.938

Table 2: The analytical data of Mb, obtained from the amino acid sequence

2. Methods

Qiagen Maxi Plasmid Purification Protocol

The bacterial pellet from an overnight eighteen hour culture was resuspended in 10 ml of solution P1, resuspension buffer (50 mM Tris-HCl, pH 8.0, 10 mM EDTA, 100 μg/ml RnaseA, 4°C). 10 ml of solution P2, lysis buffer (200 mM NaOH, 1% SDS), was added, and the solution mixed gently four to six times and incubated at room temperature for five minutes. 10 ml of chilled P3, neutralisation buffer (3.0 M potassium acetate, pH 5.5, 4°C), solution was added, the solution mixed gently four to six times and incubated on ice for twenty minutes. The solution was then centrifuged at 20,000 g for thirty minutes at 4°C and the supernatant promptly removed. The supernatant was then centrifuged for a further fifteen minutes under the same conditions.

A Qiagen tip-500 was equilibrated by applying 10 ml of buffer QBT, equilibration buffer (450 mM NaCl, 50 mM MOPS, pH 7.0, 15% isopropanol, 0.15% Triton X-100), and the column allowed to empty under gravity. The supernatant resulting from step five was applied to the tip, and allowed to enter the column by gravity flow. The Qiagen tip-500 was then washed twice with 30 ml of buffer QC, wash buffer (1.0 M NaCl, 50 mM MOPS, pH 7.0, 15% isopropanol). The DNA was eluted by the addition of 15 ml of buffer QF, elution buffer (1.25 M NaCl, 50 mM Tris-HCl, pH 8.5, 15% isopropanol). The DNA was then precipitated by the addition of 0.7 volumes of room temperature isopropanol. The solution was immediately centrifuged at 15,000 g for thirty minutes at 4°C, and the supernatant removed. The DNA pellet

was washed with 5 ml of 70% ethanol, air dried for five minutes, and redissolved in 100 µl of TE (10 mM Tris-HCl, pH 8.0, 1 mM EDTA) and RNase.

Transformation of DNA

In a laboratory environment plasmid DNA can be introduced into bacteria by the artificial process of transformation. It is in such bacteria that the DNA can grow and divide. Using sterile equipment, 20 µl of a suspension of competent cells of the type BL21(DE3)pLysS were transferred to a sterile microfuge tube on ice. 0.5 µl of DNA was added and gently mixed. A blank and a test plasmid were also prepared. The microfuge tubes were then stored on ice for 30 minutes. Subsequently the tubes were heated at 42°C for exactly 42 seconds to heat shock the cells, and then placed on ice to allow the cells to chill for 2 minutes. 80 µl of SOC medium (SOC medium is obtained from Sigma, it is a complex buffered medium commonly used for growing high-efficiency competent DNA) was added to each tube and the tubes incubated at 37°C for 45 minutes to allow the bacteria to recover and to express the antibiotic resistance marker that is encoded by the plasmid. 50 µl of the cells were transferred to a Luria agar plate containing 100 µg/ml ampicillin (AMP), and these were spread over the surface using a bent sterile rod. The plates were left at room temperature until the liquid had been absorbed and then inverted and incubated at 37°C overnight.

Cultivation of protein

10 ml of Luria Broth (LB) medium containing 100 μg/ml AMP and 34 μg/ml chloramphenicol (CAP) was inoculated with a single colony from an agar plate containing the *E. coli* BL21(DE3)pLysS with GST-Rukh61. This was incubated overnight at 37°C, 200 rpm, then added to 500 ml LB similarly treated in a 2 L flask and incubated at 37°C, 200 rpm until the optical density (OD) A_{600} was approximately 0.6. Over-expression was induced by the addition of isopropyl-1-thiol-β-D-galactoside (IPTG) to a final concentration of 1 mM and incubation was continued for 180 minutes. The cells were harvested by centrifugation for 20 minutes at 8000 g, 4°C and the pellets frozen for storage.

Minimal Media

Minimal media is used to express proteins enriched with ^{13}C and ^{15}N for use in NMR studies. The media used in this work was M9 media which contains autoclaved deionized water (800 ml), 5X salts (200 ml), 1 M $MgSO_4$ (2 ml), 1 M $CaCl_2$ (0.1 ml) and filter sterilised 20% glucose (20 ml).

Preparation of Cell Extract of Rukh61

The pellets were resuspended in a solution composed of ice cold PBS (phosphate buffer solution), DNase (20 μg/ml), and $MgCl_2$ (10 mM). This was left on ice for 15 minutes while the DNA was degraded and the viscosity of the sample reduced. Dithiothreitol (DTT) (5 mM) was added to break disulphide bonds in the GST tag and to maximise GST binding to the glutathione beads. The resulting solution was sonicated on ice for 30 seconds with a 2 minute rest. This was repeated three times. The cell debris was removed through centrifugation for 15 minutes at 15000 g, 4°C and the supernatant retained.

Preparation of Cell Extracts of Mb

Sonication was used to break open the cells and to ensure that all the protein was recovered. The pellets were resuspended in ice cold STE (Sodium chloride, Tris, ETDA) buffer (40 mM Tris(hydroxymethyl)aminomethane, 150 mM NaCl, 1 mM EDTA, pH 8) with ice cold lysozyme (10 μg/ml), DNase (20 μg/ml), and $MgCl_2$ (10 mM). This was left on ice for 15 minutes and the procedure for Rukh61 followed from then on.

Preparation Of Glutathione Sepharose-4B Beads

The beads were washed twice with 10 bed volumes of phosphate buffer solution, PBS, (140 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, pH 7.4) to equilibrate them. On completion of the purification of the target proteins the beads are cleaned using a high pH buffer (0.1 M Tris-HCl, 0.5 M NaCl, pH 8.5) and a low pH buffer (0.1 M NaAc, 0.5 M NaCl, pH 4.5). This was repeated twice before storing in PBS.

Purification Using Glutathione Sepharose-4B Beads

The supernatant from the clarification of the sonication solution was filtered onto the beads, and incubated for 30 minutes at room temperature with gentle agitation. The solution was centrifuged at 500 g for 5 minutes to sediment the beads and the supernatant removed. The beads were washed with 5-10 volumes of PBS containing 2 mM PMSF (phenylmethanesulphonyl fluoride), to remove any contaminants and again centrifuged at 500 g for 5 minutes and the supernatant removed. This was repeated until the beads were clear of impurities. Once the beads appeared to be clear of impurities they were washed with a cleavage buffer (2.5 mM $CaCl_2$, 150 mM NaCl, 50 mM Tris, pH 8.0) which removed any non-specifically bound impurities that the PBS had not. The GST fusion proteins bound to the beads and were eluted using 1 volume of elution buffer (75 mM Hepes, 150 mM NaCl, 5 mM DTT, 10 mM reduced glutathione, pH 7.4). This was left to incubate for 10 minutes at room temperature with gentle agitation. The solution was centrifuged again at 500 g for 5 minutes to

sediment the beads and the supernatant retained. This was repeated until the beads were clear of proteins.

Thrombin Cleavage To Remove GST Tag

Thrombin cleavage is possible due to a 6 amino acid recognition site (Leu Val Pro Arg Gly Ser) cloned into the protein. Thrombin protease cuts between Arg and Gly.

One unit of thrombin protease is defined as the amount of thrombin required to cleave > 90% of 100 μg of a test GST-fusion protein in PBS at 22°C in 16 hours.

Determination of Protein Concentration - Dye-Binding Procedure

The Bradford assay technique was used. Bradford Reagent[36] was produced by dissolving 100 mg of Coomassie Brilliant Blue G-250 in 50 ml of 95% ethanol mixed with 100 ml of 85% w/v phosphoric acid, diluted with water to one litre and filtered. A calibration curve was produced using stocks of BSA (bovine serum albumin) of known concentration. This reagent was mixed 3 ml : 60 μl sample and the absorbance at 595 nm measured against a blank. This technique utilises the fact that the binding of Coomassie Brilliant Blue G-250 to protein causes a shift in the absorption maxima of the dye from 465 nm to 595 nm.

Affinity Chromatography - Glutathione Sepharose Purification

Affinity chromatography is an adsorption chromatographic method in which the protein to be purified is specifically, and reversibly, adsorbed by a ligand immobilised on an insoluble support matrix. Purification can be in the order of several thousand fold and sample recovery is high. Affinity chromatography also has a concentrating effect.

Sepharose is a bead-formed agarose gel which is commonly used as a matrix. It has an open-pore structure with an exclusion limit in gel filtration of 20 x 10^6 molecular weight, this makes the interior of the matrix available for ligand attachment and allows good binding capacity relative to other matrices. Sepharose 4B exhibits low non-specific adsorption, which leads to good separation. The ligand used in affinity chromatography has to exhibit specific, reversible binding affinity for the substance to be purified and it has to be attached to the matrix without destroying the binding capacity. A spacer arm can be added between the ligand and the matrix to facilitate binding of substances otherwise affected by steric interference.

The purification of recombinant fusion proteins that contain at the carboxy terminus a GST tag is achieved using an immobilised glutathione chelating group. Glutathione Sepharose 4B, available from Pharmacia Biotech, has glutathione attached via a 10 carbon spacer arm to the oxirane group of activated Sepharose 4B. The structure of

glutathione is complimentary to the GST binding site, hence purification of the GST fusion proteins is possible via affinity chromatography. Cleavage of the target protein is via a specific protease cleavage site. In the case of proteins in the pGEX-4T1 vector the GST can be cleaved from the target protein using thrombin.

Removal of Salts and Exchange of Buffer by Dialysis

The protein solution is placed in tubing of semi-permeable membrane and the dialysis tubing is placed in the required buffer. Small molecules can pass freely across the membrane whilst large molecules are retained. The membrane is made of cellulose acetate with pores of between 1-20 nm in diameter, the pore size determining the minimum molecular weight of molecules retained by the membrane. The pore size is chosen so that molecules that are larger than the pores cannot pass through the membrane, thus salt exchange is possible but protein of a molecular weight higher than the cut-off will not pass through the membrane.

Pre-treatment of dialysis tubing ensures a uniform pore size and removal of any heavy metal contaminants. When benzolylated dialysis tubing with a cut-off of 5 kDa is to be used, the tubing is prepared by soaking in water for three hours with fresh water every hour. Dialysis tubing with a higher cut-off point is treated differently, as it needs to have heavy metal contaminants removed. This type of tubing is boiled for 10 minutes in a large volume of 2% (w/v) sodium bicarbonate, 1 mM EDTA, pH 8.0 solution. The tubing is then rinsed in distilled water and boiled for 10 minutes in 1 mM EDTA, pH 8.0 solution. The tubing is cooled and stored in 20% ethanol at 4°C, it

is rinsed with distilled water before use. The tubing is handled with gloves following the first stage of the treatment.

Equilibrium between the solution inside the tubing and the solution outside of the tubing is usually reached after 3 hours with efficient stirring using 15 kDa cut-off membranes.[36] As the molecular weight cut-off decreases, the time to equilibrium increases. Thus for tubing with 5 kDa a significantly longer time is required before equilibrium is reached than for tubing with a 15 kDa cut-off.

Gel Permeation Chromatography

This technique can be used to exchange the buffer of a solution and to separate different sizes of proteins. The volume of solution that can be exchanged into a different buffer by this technique is dependent upon the size of column used. The maximum volume should not be more than 30% of the column volume, if it is more then resolution between the protein and the salt from the previous buffer will not occur. This method results in dilution of the sample.

One method of protein concentration is to exchange the buffer solution to ammonium bicarbonate. This is a volatile solution and can be removed through freeze drying. There are two methods of exchange: the first is dialysis and the second is gel permeation chromatography. The gel permeation chromatography is faster and uses less buffer to carry out. The recovery of the protein is also higher.

Sodium Dodecyl Sulphate-Polyacrylamide Gel Electrophoresis (SDS-PAGE)

Polyacrylamide gels are formed by the copolymerization of acrylamide monomers with a cross-linking agent to form a 3D lattice. The best and commonest cross-linker for PAGE is N, N'-methylene bisacrylamide. Polymerisation is initiated with ammonium persulphate and N, N, N', N'-tetramethylethylenediamine. The concentration of acrylamide in the gel determines the effective separation range of the SDS-PAGE, for example, a 10% gel separates proteins in the molecular weight range of 15-200 kDa. Some small polypeptides and peptides (< 5 kDa) cannot be completely separated as the small molecules form SDS-protein complexes of the same dimension and charge leading to them migrating together during SDS-PAGE.

SDS is an anionic denaturing detergent and a solubilizing agent. It is used with β-mercaptoethanol to dissociate proteins in a sample into their polypeptide subunits. Each subunit binds the same amount of SDS per unit weight (1.4 g SDS per 1 g protein) and takes on the same net negative charge. Thus, all peptides in the sample will have the same charge density and separation will occur based on molecular weight. When electrophoresis begins the proteins migrate through the gel, separating according to size. This type of gel is used to determine the sizes of the proteins present in a sample and the purity.

The gels are stained by Silver or Coomassie Brilliant Blue. Silver staining is a more sensitive technique and is used for gels run on the Phast^{TM} system, where typically, samples contain only 5 ng of protein and both the separation and staining procedures are automated. The gels for this system are pre-poured by the manufacturer.

Coomassie Brilliant Blue staining of gels prepared in the laboratory involves leaving the gels with gentle agitation in a large excess of Coomassie Brilliant Blue for 4 hours. The excess dye is removed by draining off the solution and rinsing the gel with an excess of a solution of 45% MeOH, 10% acetic acid.

Silver staining of a gel not run on the Phast™ system is carried out as follows. All of the rinses are with gentle agitation. The gel is first rinsed for 15 minutes in a solution of 50% MeOH, 10% acetic acid. The next rinse is 15 minutes in 5% MeOH, 7% acetic acid. These two rinses prefix the gel, final fixing is achieved using 10% gluteraldehyde solution for 15 minutes. The gel is then rinsed in distilled water for 2 hours with fresh water after 1 hour. The gel is next rinsed in 5 μg/ml DTT for 15 minutes and then in 0.1% silver nitrate for 15 minutes. The gel is then ready to be developed using 3% sodium carbonate, 0.02% formaldehyde solution. The gel is darkened until it is the required colour and the reaction stopped with 5% acetic acid.

For a gel run on the Phast™ system the silver staining technique is similar but the system is automated and thus the time for development of the gel is pre-set. The first step is a wash in 50% EtOH, 10% acetic acid. The next wash is in 10% EtOH, 5% acetic acid. Following this is a sensitising step using 8.3% gluteraldehyde in water. The gel is then washed in deionized water. Following this the gel is treated with 0.25% (w/v) silver nitrate and then developed using 0.015% formaldehyde in 2.5% sodium carbonate. The development stage is stopped using 5% acetic acid in water. As a final stage the gel is treated with preserving solution of 10% acetic acid, 10% glycerol.

4. Results And Discussion

4.1 Expression And Purification Of Rukh61

Expression of Rukh61

Rukh61 was obtained in the form of DNA which was then transformed into the *E. coli* strain BL21(DE3)pLysS using the protocol described earlier.

GST-fused Rukh61 has been over-expressed in quantities up to 25 mg/L of pure protein. The pLysS cells in which Rukh61 has been expressed releases lysozyme, which breaks the cells open on freeze-thawing. Sonication can be used for the same effect and was used to ensure that all the protein was recovered. It was discovered during the course of this work that proteolytic cleavage was occurring. To prevent the Rukh61 degrading through this proteolysis phenylmethanesulphonyl fluoride (PMSF), from a 0.2 M stock in ethanol, was added to all of the solutions except those in which thrombin cleavage occurs. When PMSF is only added to the lysis buffer some degradation of GST-Rukh61 still occurs during the purification process. PMSF is used to inactivate serine proteases. Cysteine proteases can be inactivated through the use of N-ethyl maleimide (NEM) which is added at lysis only. This was included in the protocol because a band at 20 kDa appeared on gels after lysis that was not present before. The addition of NEM solved this problem.

Purification of GST-Rukh61

At the start of this project sarkosyl (1.5%) was added to the lysis solution to

encourage the solubilization of the protein. It was found that this was not required as GST-Rukh61 is expressed in a soluble form. During this time Triton X-100 (1%) was added to encourage binding to the glutathione Sepharose beads, as sarkosyl solubilized GST does not bind well to agarose beads without the use of a detergent.

After lysis the solution was clarified by centrifugation. The clarified protein solution was applied to the glutathione Sepharose 4B beads and the impurities removed by washing with PBS containing 2 mM PMSF. The Bradford Assay was used on the wash fractions to determine if all of the impurities had been removed. Table 3 shows a typical set of Bradford readings for the cleaning of the beads and Figure 7 shows a gel of the fractions.

	mg/ml of protein
Pass (20ml)	1.243
Wash 1 (40ml)	0.842
Wash 2 (40ml)	0.373
Wash 3 (40ml)	0.156
Wash 4 (40ml)	0.094
Wash 5 (40ml)	0.065
Wash 6 (40ml)	0.010

Table 3: The Bradford readings from a typical purification of Rukh61.

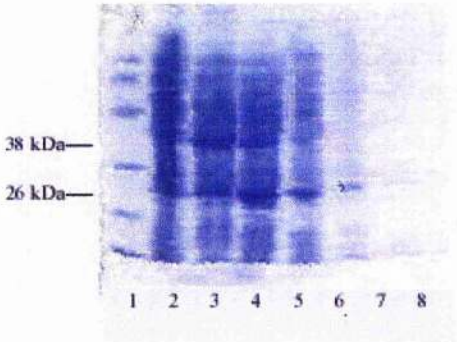


Figure 7: Gel of the stages of washing of the glutathione beads.
This is a 12% acrylamide gel that was stained using Coomassie Blue. The markers in lane 1 are at 94 kDa, 67 kDa, 43 kDa, 30 kDa, 20.1 kDa and 14.4 kDa (largest at the top). Lane 2 is the pass of supernatant from the clarification of the beads. Lanes 3-8 are washes 2-6.

Using the markers in lane 1 of Figure 7 as a reference, it can be seen that there is a band at 38 kDa and at 26 kDa in lanes 2–5. As lane 2 is the supernatant and lanes 3-8

are washes of the beads, this implies that GST-Rukh61 and GST have both been expressed but have not bound completely to the beads. This problem can be solved by the addition of Triton X-100 to the supernatant before it is applied to the beads. This can lead in turn to the creation of another problem, the concentration of Triton X-100 in the supernatant must not be so high as to cause unfolding of the protein. If Triton X-100 is not added to the solution from sonication then the target protein does not bind as well to the beads. If it is found that not all of the protein has bound to the beads then the solution is reloaded and the remaining protein purified. It was found that Triton X-100 at 0.1% enables binding to the glutathione beads to be increased. It was also found that if the beads are used repeatedly then the binding capacity deteriorates.

Figures 8 and 9 show the elution of GST-Rukh61 from the glutathione Sepharose beads. These gels show that in Rukh61 there is a pseudo thrombin site, this is GKA and the thrombin is able to cleave between the K and A. When the thrombin protease was allowed to work on the protein for longer than 20 hours then cleavage at this pseudo site occurred. Cleavage at this site gave two smaller proteins rather than Rukh61. Table 4 shows basic data that was obtained about these smaller proteins using the computer program 'Protean'.[20]

Molecular weight	3.816 kDa	8.097 kDa
pI	9.007	9.080
Charge at pH 7.0	0.911	1.926

Table 4: Proteins produced by cleavage at pseudo thrombin site.
The data in this table was obtained using 'Protean', an analysis program, and is based on the amino acid sequence of Rukh61 cleaved at the pseudo thrombin site.

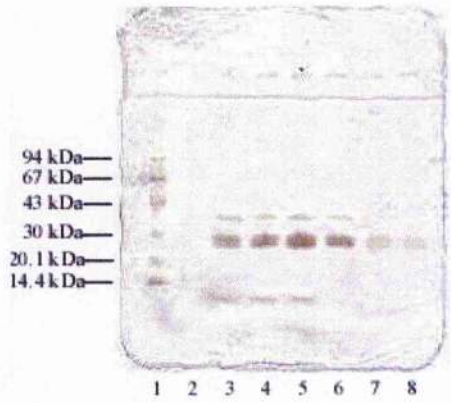


Figure 8: The elution of GST-Rukh61 from glutathione Sepharose beads when PMSF is not used.
Lane 1 contains the markers at 94 kDa, 63 kDa, 43 kDa, 30 kDa, 20 kDa and 14 kDa (largest at the top). Lane 2 is empty. Lanes 3-8 show elution fractions.

Figure 8 shows the elution fractions obtained from glutathione Sepharose beads when PMSF is not used in the purification procedure. Lanes 3-5 show the fused protein at 38 kDa, GST at 26 kDa and a band that appears to be cleaved Rukh61 at 11 kDa (faint line). Lane 6 shows the fused protein at 38 kDa and GST at 26 kDa. Lanes 7, 8 show only GST. This gel shows that proteolytic cleavage occurs when PMSF was not used in the elution and washing buffers. The bands at 26 kDa appear to be double bands, this implies that cleavage was occurring at more than one site. The proteases normally present should not be cleaving at the thrombin protease cleavage site but the GST-Rukh61 is cleaving in such a way as to suggest that this is occurring.

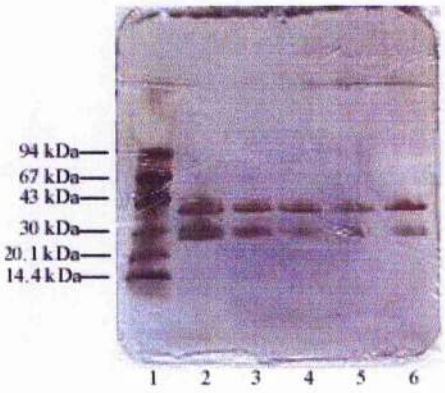


Figure 9: The elution of GST-Rukh61 from glutathione Sepharose beads when PMSF has been included in the buffers.
Lane 1 contains the markers at 94 kDa, 63 kDa, 43 kDa, 30 kDa, 20 kDa and 14 kDa (largest at the top). Lanes 2-6 show elution fractions.

Figure 9 shows the elution fractions obtained from glutathione Sepharose beads when

PMSF is used in the purification procedure. In contrast to Figure 8, there are only bands at 38 kDa (fused protein) and at 26 kDa. This implies that the use of PMSF prevents protease cleavage occurring. The presence of a 26 kDa band in the elution fractions of GST-Rukh61 implies that the fused protein is cleaving whilst expression is occurring. The band at 26 kDa is not likely to be an *E. coli* protein.

The elution buffer used to remove the GST-fused protein from the glutathione beads contained reduced glutathione. This had to be removed from the protein solution prior to thrombin cleavage as it interferes with the reaction. The removal can be carried out through dialysis against PBS, however as the glutathione was removed, the protein fell out of solution. Thus, at a concentration of only 1 mg/ml GST-Rukh61 became insoluble. Altering the pH to 6.5 improved the solubility, as did reducing the salt content. Eluted GST-fused Rukh61 can be concentrated using dialysis, however, concentration using centrifugation caused cleavage of the protein to occur through shearing against the membrane.

Thrombin Cleavage

Time course experiments at different temperatures have shown that the temperature for fast cleavage of the GST tag from Rukh61 is 37°C with a gentle agitation of 75 rpm but at this temperature cleavage also occurs immediately at the pseudo site. Time studies of the cleavage have also been carried out and these show that cleavage overnight (20 hours) at 22°C (room temperature) is enough to cut all of the protein present if sufficient thrombin protease is present in the solution. The use of 2 units of thrombin protease per mg of protein is sufficient, and does not lead to excessive

cleavage at the pseudo thrombin site.

Prior to thrombin cleavage the GST-fused Rukh61 was originally dialysed into PBS but method development has lead to the use of a buffer known as the cleavage buffer (see Section 2 for details). The cleavage was carried out using 2 units of thrombin per mg of GST-Rukh61. Thrombin cleavage yields 5 mg of cleaved protein from 18 mg of uncleaved protein produced by 1 L of broth.

The cleaved protein solution is re-applied to the glutathione beads to remove the GST from the solution. This is because the GST binds to the beads but the Rukh61 does not.

Thrombin Cleavage on Glutathione Sepharose Beads

An alternative method of thrombin protease cleavage which eliminates glutathione totally from the protein solution is to cleave the GST-Rukh61 whilst it is attached to the Glutathione Sepharose 4B beads. In this case the elution step of the purification protocol was not carried out, instead 1 volume of cleavage buffer was applied to the beads after the washes and thrombin protease added to a final concentration of 2 units/mg of fused protein.

The amount of fused protein present can be estimated from prior knowledge of yields. Determination of the yield of GST-Rukh61 prior to cleavage is carried out by the removal of 0.5 ml of the beads and eluting it with elution buffer. The total yield was then calculated from the yield of that fraction.

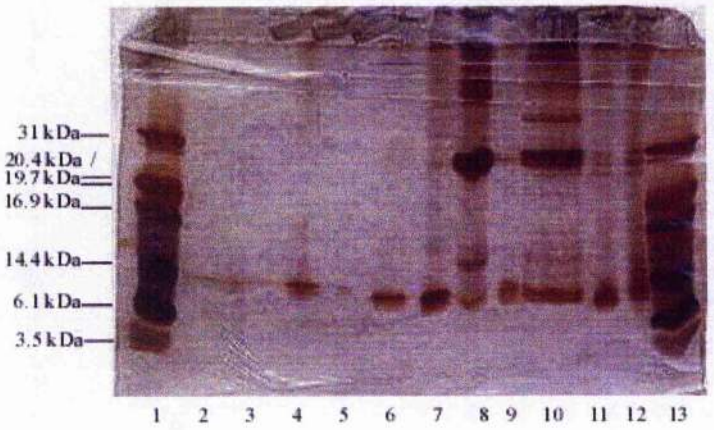


Figure 10: Gel of the cleavage of GST-Rukh61. This is a 15% SDS-PAGE acrylamide gel that has been silver stained. Lanes 1 and 13 contain low molecular weight (LMW) markers, as shown. Lanes 2-10 show stages of cleavage. Lanes 11-12 show how purity can vary with different wash buffers.

To analyse the cleavage of the GST-Rukh61 samples were taken after ½ hour, 1 hour, 2 hours, 3 hours, 5 hours and 20 hours. These are shown in Figure 10, lanes 2-7. After 6 hours of cleavage the beads were split into two batches: one was left overnight until it had been reacting for 20 hours and the other batch was eluted immediately. Lane 8 is the elution of the beads after 20 hours of cleavage, showing GST and Rukh61 only. Lane 9 shows a sample taken after 6 hours and lane 10 shows the elution after 6 hours, with fused Rukh61, at 38 kDa still present. This gel also shows cleavage occurring at the pseudo thrombin site, which can be seen by the presence of a double band at 11 kDa. Lane 11 is an example of cleavage of a sample cleaned with cleavage buffer and lane 12 is an example of cleavage when a wash with cleavage buffer has not been carried out.

Concentration of the protein

Ultrafiltration

The most commonly used method of concentration is the use of ultrafiltration in an Amicon concentrator where pressure and a stirring bar are used to force the solvent through a membrane. Initially protein is lost to the membrane using this method

however, once the membrane has been saturated with the required protein, no more is lost. The concentration of Rukh61 was carried out using the common method of ultrafiltration in an Amicon concentrator. This method can have a high percentage of loss for small amounts of proteins. For example, a sample containing 11 mg of protein was concentrated this way and 70% of the protein was lost, it did not pass through the membrane and soaking the membrane to remove loose protein did not increase the yield. This problem may be solvable through saturating the membrane with another protein or through changing the membrane type. If the protein is unfolded then it may be sticking to the container in which case the above solutions would have no effect.

Centrifugation

Centrifugation methods can be used as the membranes can be treated with glycerol to prevent the proteins sticking, unfortunately centrifugation methods cannot be used for the concentration of GST-Rukh61 because the protein appears to be sheared using this method.

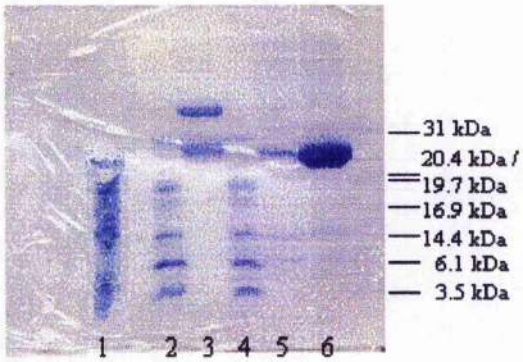


Figure 11: Gel of the products from centrifugation of GST-Rukh61.
Lane 1 contains a sample from the waste solvent side of the membrane. Lanes 2 and 4 contain LMW markers. Lane 3 contains the original sample of GST-Rukh61. (Lanes 5 and 6 contain samples of GST eluted from a column after cleavage of the Rukh61)

The gel in Figure 11 shows the products of centrifugation in Lane 1, these protein fragments have passed through the membrane. The sample seen in lane 3 is the original sample prior to centrifugation, it can be seen that this sample contains GST as well as GST-Rukh61. Lanes 2 and 4 contain LMW markers, the numbers shown on

the side of the gel refer to these markers. Lanes 5 and 6 contain samples of GST eluted from a column after cleavage of the Rukh61 and are not connected with centrifugation. As can be seen there is protein passing through the membrane and small proteins are visible which are not seen in the original sample. This implies that the protein is being sheared by the centrifugation concentration procedure.

Dialysis

Dialysis against a solution of 20 K PEG is one method of concentrating protein or the use of dry PEG for a quicker concentration. These both lead to a loss of GST-Rukh61 and of Rukh61 and to the protein falling out of solution when it reaches 1 mg/ml. Dialysis into ammonium bicarbonate with a view to freeze drying leads to a 20% loss of the protein when there is initially 8 mg of protein.

Gel permeation chromatography

The concentration of protein can be achieved through the exchange of protein into a suitable solution such as ammonium bicarbonate and freeze drying to remove the volatile solvent. Unfortunately concentration through the use of a G25 column and freeze-drying can lead to unfolding of proteins. Further investigations into this technique can be carried out once it is certain that the protein is folded to start with.

The gel in Figure 12 shows the fractions from a G25 Sephadex column after freeze-drying. Lane 1 contains LMW markers. Lanes 2-6 contain fractions 8-12 of the 5 ml fractions collected. Lane 7 shows a sample containing Rukh61 and GST.

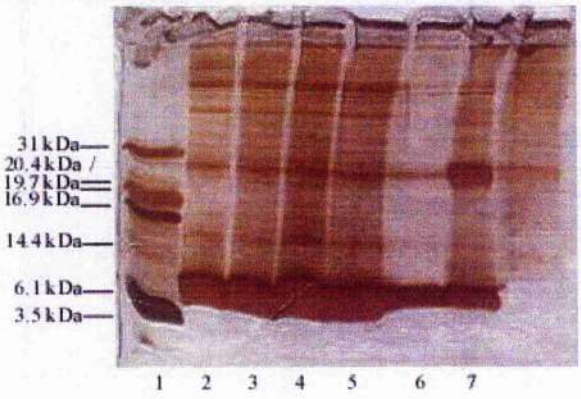


Figure 12: Fractions from G25 Sephadex column.
Silver stained 15% acrylamide gel of the fractions from a G25 Sephadex column after freeze-drying. Lane 1 contains LMW markers. Lanes 2-6 contain fractions 8-12 of the 5 ml fractions collected. Lane 7 shows a sample containing Rukh61 and GST.

The gel in Figure 12 was overloaded to show up the impurities which are from non-specific proteins binding to the glutathione beads. These impurities can be removed from the beads by washing them with the cleavage buffer prior to the cleavage step. The purity of the samples obtained by washing with cleavage buffer is shown in Figure 13.

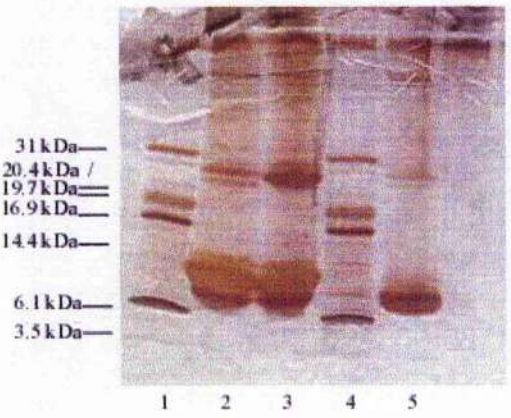


Figure 13: Cleavage of Rukh61-GST.
Lane 1 contains LMW markers as does lane 4. Lane 2 shows the supernatant and lane 3 shows the beads after 20 hours of cleavage. Lane 5 shows the centrifuged supernatant, containing cleaved Rukh61.

In Figure 13, lane 1 contains LMW markers as does lane 4. Lane 2 shows the supernatant and lane 3 shows the beads after 20 hours of cleavage. Lane 5 shows the centrifuged supernatant, containing cleaved Rukh61. These samples are a lot purer than those where PBS alone has been used, as in Figure 14. This is the method that is now used.

It has been found that the cleavage buffer removes the majority of these non-specifically bound impurities from the beads, leading to a more pure product after thrombin cleavage, prior to application to the G25 Sephadex column.

Ammonium sulphate precipitation

Another method of concentration of a protein is ammonium sulphate precipitation where the protein is precipitated by the addition of ammonium sulphate and resuspended in a smaller volume of buffer. The addition of ammonium sulphate to 80% is required to precipitate the protein. This can then be pelleted by centrifugation at 20,000 g for 30 minutes. This method of concentration is presently being looked into, however preliminary experiments indicate that this method of concentration leads to a high percentage yield with few losses.

Table 5 shows examples of the different concentration techniques and the losses that occur on their use. The techniques that retain the highest percentage of Rukh61 are, unfortunately, those that are most damaging to the protein.

Concentration Technique	% Rukh61 remaining
G25 column	97
Dialysis against ammonium bicarbonate	78
Dialysis against PEG	32
Amicon concentrator	30

Table 5: Different concentration techniques that were tested and the amounts of Rukh61 remaining at the end of the concentration procedure.

Growth of Rukh61 under differing conditions

GST-Rukh61 has been over-expressed in Luria broth, Terrific broth and in M9 media before purification on the glutathione Sepharose 4B beads. The yields for these different media are shown in Tables 6-8 below. Rukh61 was expressed in Terrific broth because it has previously lead to higher yields, in this case it did not. It was expressed in M9 (minimal) media as this is the method used to produce labelled protein for NMR studies.

Tables 6-8 show the results of these experiments in different media. It can be seen from the data that the yield of GST-Rukh61 is similar regardless of the media in which it is produced. GST-Rukh61 was expressed in Luria media on six occasions to provide the data in the table below. It was expressed once in M9 media and in Terrific Broth with the intention to repeat the work when other problems had been solved and the method of expression had been further optimised.

Luria Broth	Yield (mg/L)	% Total Protein	Error (mg)
Total Protein	110 mg		± 20
Pure GST-Rukh61	25 mg	22.7	± 4
Cleaved Rukh61	7 mg	6.4	± 0.5

Table 6: The expression of GST-Rukh61 in Luria broth.
This shows the percentage of total protein obtained in the form of Rukh61. The experiment was repeated six times to calculate the expected yields and errors.

Terrific Broth	Yield (mg/L)	% Total Protein
Total Protein	91 mg	
Pure GST-Rukh61	15 mg	16.5
Cleaved Rukh61	4 mg	4.4

Table 7: The expression of GST-Rukh61 in Terrific broth.
This shows the percentage of total protein obtained in the form of Rukh61.

M9 Media	Yield (mg/L)	% Total Protein
Total Protein	88 mg	
Pure GST-Rukh61	14 mg	15.9
Cleaved Rukh61	3.5 mg	4.0

Table 8: The expression of GST-Rukh61 in M9 media.
This shows the percentage of total protein obtained in the form of Rukh61.

Due to difficulties in the resuspension of GST-Rukh61 from pellet form, possibly due to the formation of inclusion bodies, attempts have been made to express it at different temperatures and shaker speeds, standard LB being used, in the hope that this would solve these problems. The results of these experiments are shown in Table 9 below, and from these results it can be seen that at 30°C and at 37°C, 100 rpm the growth was minimal. The use of lower temperatures and/or shaker speeds could help with the folding of the protein too.

Temp (°C)	Speed (rpm)	Total protein (mg/L)	Total GST-Rukh61 (mg/L)	% GST-Rukh61
37	200	110	25	22.7
37	100	42	0.8	1.9
30	200	62	1.1	1.8

Table 9: The production of GST-Rukh61 under differing experimental conditions.
All of the experiments were carried out in Luria broth. The experiments not at 37°C and 200 rpm were repeated twice in order to check that the results were standard.

Overview

There are certain problems which have arisen in the course of this work which point toward the possibility that GST-Rukh61 (and therefore Rukh61 also) was incorrectly, or not at all, folded. The first was that it did not bind as well to the glutathione

Sepharose 4B beads as was expected, there should have been no need to re-apply the supernatant to the beads to recover further protein. The second indication was that Rukh61 and GST-Rukh61 was very difficult to concentrate, once the concentration reached 1 mg/ml the remaining protein precipitated. In addition to this, protein was lost on concentration and dialysis through sticking to the membranes and to the sides of the container. This would not occur in a folded structure because in a coiled-coil the hydrophobic groups are in the centre of the molecule.

A third pointer was that NMR on a sample of GST-Rukh61 failed to show any sign of folding that would normally be seen even with a weak sample. See Appendix 2 for the spectra.

A sample containing folded protein would have many more peaks between 1-2 ppm than are seen in this sample. The peaks that are seen at 1-2 ppm in a folded sample arise from through space interactions rather than through bond interactions. The peaks observed in this NMR spectrum show the groups present in the proteins, CH_2, OH, NH_2 etc. but in a 2/3D NMR spectrum this particular sample would not give information on GST-Rukh61 conformation.

4.2 Expression And Purification Of Mb

The Mb was originally received in the *E. coli* strain X90 but the level of expression of the target protein was low in comparison to the overall expression. Prior to transformation into another strain, DNA had to be prepared. A single colony is grown to saturation in 5 ml of media then diluted 1:100 and again grown to saturation. The DNA is then isolated using the QIAGEN protocol for plasmid purification as shown in Section 2.

Expression of GST-Mb

From the gel in Figure 14 we know that GST-Mb has been produced. When STE and lysozyme is used as the resuspension buffer the pellet is easier to resuspend and does not form clumps as it does when PBS is used. Purification on glutathione Sepharose 4B beads has yielded 5 mg/L of protein, but it appears that a 30 kDa protein is the major product.

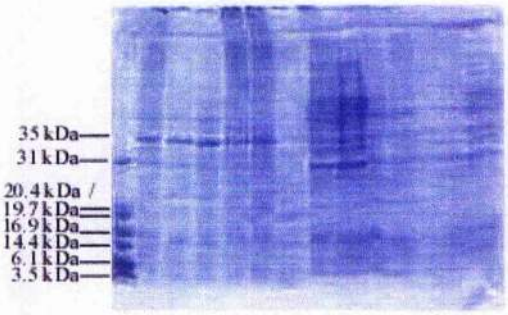


Figure 14: Gel of the expression of GST-Mb.
Lane 1 contains low molecular weight (LMW) markers, 31 kDa, 20.4 kDa, 19.7 kDa, 16.9 kDa, 14.4 kDa, 6.1 kDa, and 3.5 kDa (from the top to bottom). Lanes 2–14 show the growth of GST-Mb after induction with IPTG, with sampling at 30 minute intervals.

Figure 14 shows the expression of GST-Mb. Lane 1 contains LMW markers, at the weights shown in the diagram. Lanes 2–14 show the growth of GST-Mb after induction with IPTG, with sampling at 30 minute intervals. As seen in the figure, the

30 kDa protein mentioned above can be seen two hours after induction. At this point the GST-Mb, shown as a band at 35 kDa, began to deteriorate and the unknown band at 30 kDa began to be produced. This led to lower yields of GST-Mb than expected with a high yield vector. In an attempt to overcome this problem of deterioration, the method of production of GST-Mb was altered so that the cells grow until OD_{600}=1, induction with IPTG is carried out, and the cells are then harvested after 2 hours. The concentration of IPTG used to induce growth was altered, from 1 mM to 2 mM. This led to a higher yield of GST but not of GST-Mb which appeared to still be deteriorating. This is shown by the data in Table 10 and Table 11. These tables show how the GST-Mb appears to deteriorate, so that whilst with 2 mM IPTG there is a larger amount of total protein produced, less of it is GST-Mb.

Luria Broth 1 mM IPTG	**Yield (mg/L)**	**% Total Protein**	**Error (mg)**
Total Protein	118 mg		± 20
Pure GST-Mb	7 mg	5.9	± 3
Cleaved Mb	1.5 mg	1.3	± 0.5

Table 10: The expression of GST-Mb in Luria broth with 1 mM IPTG.
This shows the percentage of total protein obtained in the form of Mb. The growth was repeated four times to give an estimate of the error.

Luria Broth 2 mM IPTG	**Yield (mg/L)**	**% Total Protein**	**Error (mg)**
Total Protein	152 mg		± 26
Pure GST-Mb	5 mg	3.3	± 2
Cleaved Mb	1 mg	0.7	± 0.5

Table 11: The expression of GST-Mb in Luria broth with 2 mM IPTG.
This shows the percentage of total protein obtained in the form of Mb. The growth was repeated twice times to give an estimate of the error.

M9 Media	Yield (mg/L)	% Total Protein
Total Protein	50 mg	
Pure GST-Mb	1.7 mg	3.4

Table 12: The expression of GST-Mb in M9 media.
This shows the percentage of total protein obtained in the form of GST-Mb. This growth was not repeated, nor was the GST-Mb cleaved as it was believed that further work on the optimisation of the procedure was required first.

Table 12 shows the one experiment carried out with M9 media and GST-Mb. The amount of GST-Mb produced led to the decision to work further on optimisation of the protocol before further minimal media experimentation.

Elution and cleavage of GST-Mb

The elution buffer used to remove the GST-fused protein from the glutathione beads contains reduced glutathione. This had to be removed from the protein solution prior to thrombin cleavage as it interferes with the reaction. The removal can be carried out through dialysis against PBS, but as the glutathione was removed, the protein was found to fall out of solution. This could be caused by the protein not being properly folded, that is, at a concentration of only 1 mg/ml, GST-Mb became insoluble. The addition of a reducing agent, $Na_2S_2O_4$, improved the solubility to 1.5 mg/ml, as did reducing the salt content and altering the pH to 6.5. GST-fused Mb could then be concentrated using centrifugation or dialysis.

Thrombin cleavage was carried out using 2 units of thrombin per mg of GST-Mb, where a unit is defined as the amount of thrombin protease required to cleave > 90% of 100 μg of a test GST-fusion protein in PBS at 22°C in 16 hours. Thrombin cleavage yielded 1 mg of cleaved protein from 5 mg of uncleaved protein.

The cleaved protein solution was applied to the glutathione beads to remove the GST from the solution. The GST binds to the beads, the Mb does not.

Concentration of the protein

The concentration of Mb can be carried out using the common method of ultrafiltration in an Amicon concentrator where pressure and a stirring bar are used to force the solvent through a membrane. This technique has led to a high percentage loss when the original amount of protein is low. Centrifugation through a semi-permeable membrane can also be used, but has not been attempted to date. When the concentration of Mb reached 1 mg/ml the protein began to precipitate. This problem is still under consideration, however Mb appears to be more soluble in a buffer containing less salt than there is in PBS and containing a reducing agent such as $Na_2S_2O_4$.

Overview

GST-Mb has been expressed in the crude extract but there is a lot more to be done before structural studies can begin.

5. Conclusion and Further Work

The aim of this research was to produce and purify the proteins Mb and Rukh61, and to determine their 3D structure by NMR. Rukh61 has been expressed but the quantities of pure protein that were being produced were small. Further investigations into the conditions of production may lead to a higher yield. A variety of problems arose in the course of the purification of Rukh61. The proteolytic cleavage and non-specific binding both led to errors in yield calculations and problems with purity but it is thought that both of these problems are now solved. The concentration of Rukh61 is still under investigation, as is the solubility in a variety of buffers. Once a suitable quantity of Rukh61 has been produced it can be dissolved in a buffer suitable for NMR, that is, a buffer which does not interfere with the NMR spectrum such as a phosphate buffer.

Mb has only been produced in very small quantities as yet. Work on Mb was put off until NMR collection on Rukh61 had been started. This unfortunately has not been achieved. Further work on Mb would have begun with the optimisation of the expression and purification of the GST-Mb and of the cleavage of the GST tag from the target protein.

The first NMR techniques which would have been used are COSY, NOESY and TOCSY. The COSY spectra will show the ^{1}H-^{1}H through bond coupling. The NOESY spectra will show the through-space distances between the protons, and the TOCSY will show the connectivities within each amino acid. Later, more complex experiments would have been undertaken and those spectra would have been analysed using data from the earlier experiments.

6. Appendix

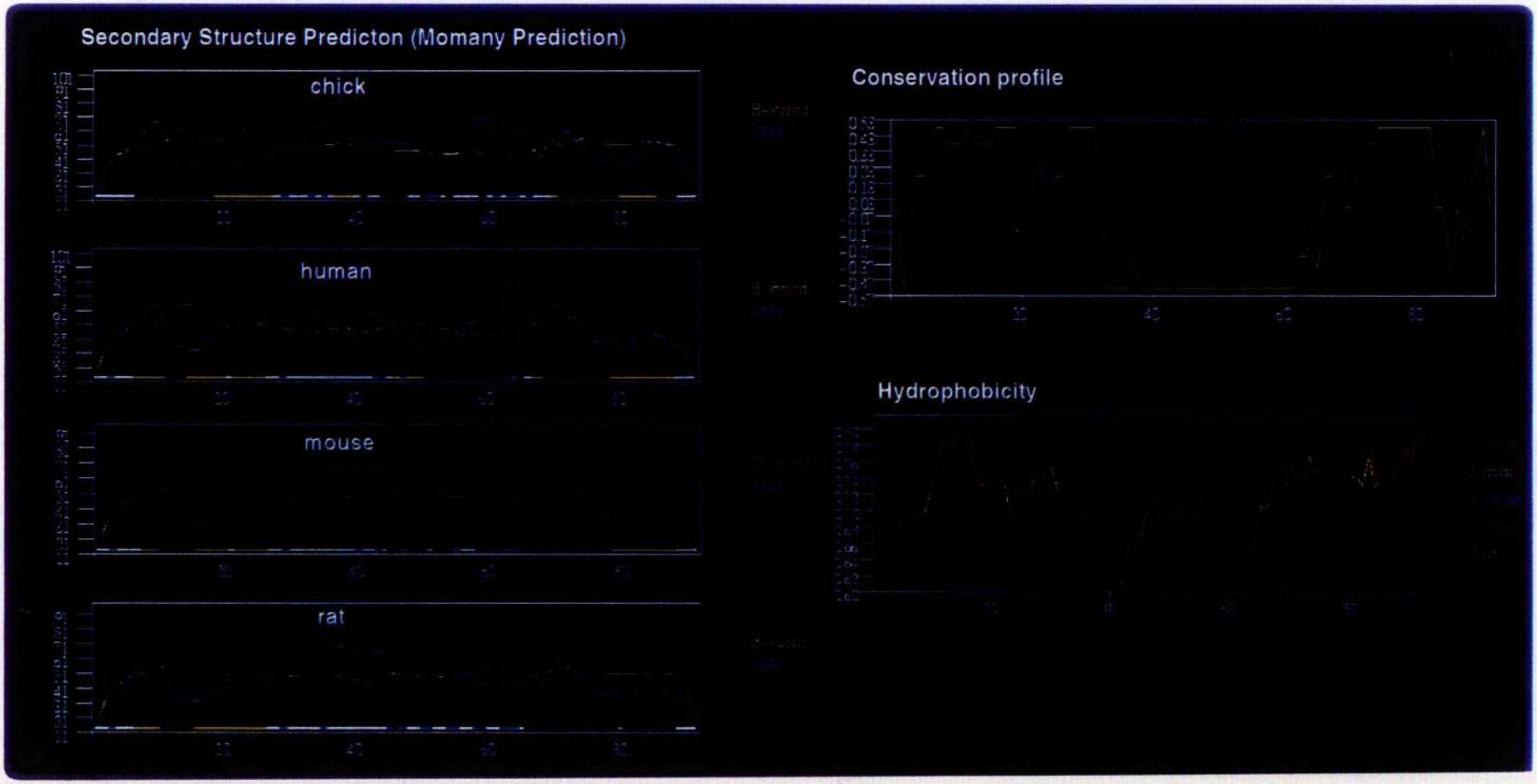
6.1 Structure Prediction For Mb

6.2 NMR Spectra of GST-Rukh61

6.3 NMR Spectroscopy

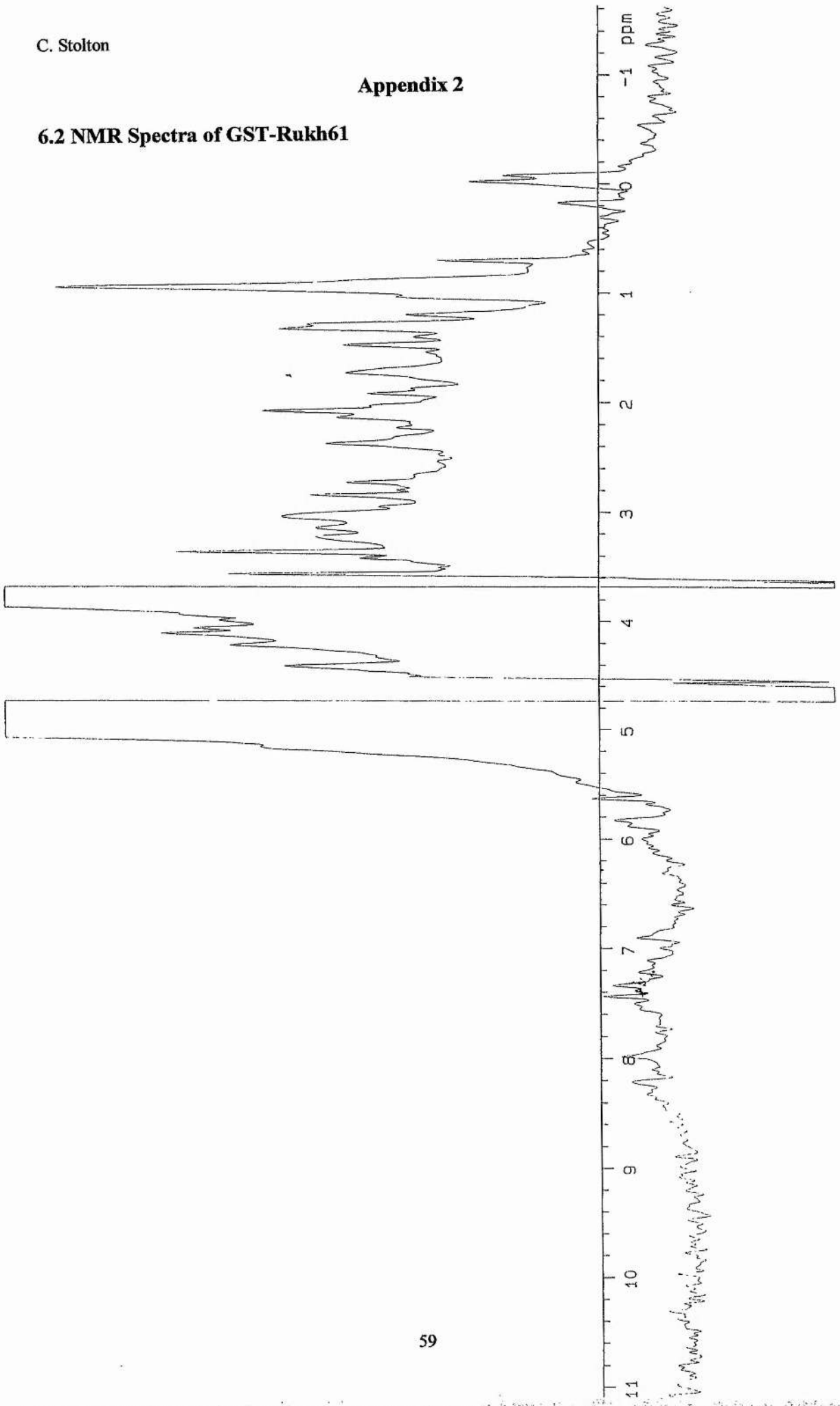
```
if_human   *1  *mrdpvssqys*11 *sflfwrmpip*21 *eldlselegl*31 *glsdtatykv*41 *kds svgkmi*50 *gqataadqek
if_mouse   *1  *mrdpvssqys*11 *sflfwrmpip*21 *eldlselegl*31 *glsdtptyes*41 *kdssvgkmn*51 *gqasgteq k
if_rat     *1  *mrdpvssqys*11 *sflfwrmpip*21 *eldlselegl*31 *glsdsptyks*41 *kesnsigkmg*51 *gqatgaer k

if_chick   *53 *caeedsllpf*63 *ssfnfwrapi*73 *asissfdfdl*83 *i
if_human   *60 *npegdglley*70 *stfnfwrapi*80 *asihsfeldl*90 *l
if_mouse   *60 *npegdplley*70 *stfnfwrapi*80 *asihsvdldl*90 *l
if_rat     *60 *spegdplley*70 *stfnfwrapi*80 *asirsidldl*90 *l
```



Appendix 2

6.2 NMR Spectra of GST-Rukh61



Appendix 3

6.3 NMR Spectroscopy

Nuclear magnetic resonance was first observed in 1946, and it is the only structural technique which can encompass three states of matter. It is the study of the properties of a molecule containing magnetic nuclei. It is a radio frequency technique which looks at the interaction of radio-frequency pulses with nuclei and the transfer of magnetisation between these nuclei.

Identifying the 3D structure of a protein is useful in understanding how that protein functions. The first step in the structure determination of proteins by NMR is the assignment of the ^{1}H, ^{13}C, ^{15}N chemical shifts, followed by the assignment of nuclear Overhauser effect (NOE) interactions. The ^{13}C and ^{15}N data is obtained from isotopically enriched proteins. The principle source of the geometric information that is used to solve 3D structures of proteins by NMR comes from the short (< 5 Å) interproton distance restraints derived from the NOE measurements. Backbone experiments[37] are used to assign the backbone atoms, that is H^{α}, C^{α}, CO, N, and NH. Side chain experiments are used to assign the resonances of the side chain atoms.

NMR is a useful analytical technique because it uses an environment close to biological conditions, it gives information on dynamics and can identify side-chain motion. NMR requires solutions concentrated to > 10 mg/ml. NMR is currently limited to proteins of < 300 (~30 kDa) amino acids due to the number of resonances in larger proteins giving an overcrowded, overlapped spectrum which is very difficult to assign, as well as transverse relaxation effects. The reliability of models derived

from NMR data is measured by the number of constraints per residue that can be derived from the experimental data. The accuracy of the NMR structure depends on the choice of method for structure calculation, restraining the molecular dynamics gives a greater accuracy than using purely distance geometry from the data.[38]

The basic concept of 2D NMR spectroscopy is to record the NMR signal as a function of time in two dimensions, the key step in 2D NMR being the transfer of magnetisation between different nuclei. There are two principle categories of 2D-NMR experiments, one is correlated spectroscopy where the cross peaks occur only between nuclei that are covalent neighbours, that is, no more than three chemical bonds apart. The other type of experiment involves NOE spectroscopy where the intensity of the cross peaks illustrates the through-space separation of the nuclei.

NMR experiments fall into two categories, homonuclear correlation experiments and heteronuclear correlation experiments. The fundamental NMR experiments are COSY (homonuclear), NOESY (homonuclear), HSQC (heteronuclear) and TOCSY (homonuclear).

COSY is a 2D technique which gives rise to diagonal and cross peaks. The cross peaks come from the transfer of magnetisation between spins. The peaks have a multiplet structure from the coupling.

TOCSY is a 2D technique which gives rise to diagonal and cross peaks. Some of the cross peaks are relay peaks arising due to the magnetisation on the protons moving to protons further away than one bond. This occurs during the spin lock sequence and

allows the connectivities of a length of C-H or N-H bonds to be traced. The TOCSY relies on the cross-polarisation caused by a spin-locking field. For a simple two-spin system complete exchange of magnetisation occurs for spin-lock mixing times of ½ J, where J is the coupling constant. For biological macromolecules involving large spin systems spectra are obtained with a variety of spin locking times.

NOESY is a 2D technique which shows cross peaks between protons that are less than 5 Å apart in space. The cross peaks are caused by the Nuclear Overhauser Enhancement which is the interaction of nuclei through space.

Polarisation transfer can be used to enhance another nucleus. The basic pulse sequence for heteronuclear correlation spectroscopy is the 'Insensitive Nuclei Enhanced by Polarisation Transfer' or INEPT. The INEPT sequence is found in many more complex experiments where it is used to transfer magnetisation from one nucleus to another heteronucleus.

HMQC was introduced on the principle that optimum sensitivity would be obtained by starting with proton polarisation and ending with proton detection, with indirect detection of the heteronucleus. The experiment correlates ^{1}H and ^{13}C resonances.

HSQC is the conventional 2D experiment for ^{1}H-^{13}C or ^{1}H-^{15}N correlation. It does not give any diagonal peaks as all of the magnetisation is transferred. The sequence includes two INEPT-type transfers to transfer magnetisation from protons to heteronuclei and the reverse which reconverts the $^{13}C/^{15}N$ magnetisation into observable ^{1}H magnetisation. This gives a greater enhancement in sensitivity than

arises from Nuclear Overhauser effects in heteronuclear correlation experiments. The spectra obtained have a pure absorptive lineshape with a linewidth narrower than with an HMQC due to it not being affected by homonuclear J coupling. The HSQC experiment is more sensitive than the HMQC. HMBC (heteronuclear multiple bond correlation) is a sensitive technique based on HSQC and HMQC for the determination of long range (2/3 bond) heteronuclear connectivities, for example, in proteins NH with $C_\alpha H$ of the next residue.

For ^{15}N NMR studies of proteins, single-quantum coherence experiments give better resolution than multiple-quantum[39], due to the dipolar broadening of ^{1}H-^{1}H being stronger than the heteronuclear broadening of ^{15}N. The obtainable resolution for ^{1}H-^{15}N shift is dependent on the pulse scheme used.

In NMR of proteins, the signal-to-noise ratio can be increased through the use of fractionally deuterated samples.[40] This is especially useful if the deuteration is combined with uniform ^{13}C, ^{15}N labelling. These samples can be prepared according to papers published by L. E. Kay and co-workers.[41,42] The gain in the signal-to-noise ratio results from the increased relaxation times of the H^N, N, C_α spins due to the fractional deuteration. Field gradients and selective radio frequency pulses are used to maintain precise control of the water magnetisation in some experiments. This technique can be used with ROESY[43], NOESY and TOCSY for the highfield NMR of biomolecules.

Triple Resonance Experiments For Backbone and Sidechain Assignments

The experiments HNCA, HN(CO)CA, HCACO, H(CA)NNH, and TOCSY-HMQC can be used for backbone assignments. The experiments HCCH-COSY and HCCH-TOCSY can be used for side chain assignments. Triple resonance experiments are useful for the backbone assignment of doubly labelled proteins, the side chain assignments of double labelled proteins, and for the more difficult side chain assignments such as H_ε, lysine, aromatics, histidine and tryptophan.

HNCA : In this experiment two cross peaks are seen for each amino acid in the protein.[44] The co-ordinates of the peaks are (Ω_{Ni}, Ω_{Ci-1}, $\Omega_H{}^N{}_i$) and (Ω_{Ni}, Ω_{Ci}, $\Omega_H{}^N{}_i$). This correlates the H_N and ^{15}N in one residue with the $^{13}C_\alpha$ of the same and the previous residue. The mixing uses $^1J_{NH}$, $^1J_{CN}$, and $^2J_{CN}$.

HN(CO)CA : In this experiment one cross peak is seen for each amino acid in the protein.[44] The co-ordinates of the peaks are (Ω_{Ni}, Ω_{Ci-1}, $\Omega_H{}^N{}_i$). This correlates the H_N and ^{15}N of one residue with the $^{13}C_\alpha$ of the previous residue. The mixing uses $^1J_{NH}$, $^1J_{CON}$, and $^1J_{COCA}$.

H(CA)NNH : The H_α (C_α)-^{15}N-NH experiment can be used to correlate the NH, ^{15}N, $C^\alpha H$ resonances.[44] This experiment can be used to show intra-residue connectivity between $C^\alpha H(i)$, $^{15}N(i)$, and NH(i) arising from transfer of magnetisation via one bond $^{13}C_\alpha$-^{15}N coupling. It can also show sequential inter-residue connectivity between $C^\alpha H_{(i-1)}$, $^{15}N_{(i)}$, and $NH_{(i)}$ via two-bond $^{13}C_\alpha$-^{15}N coupling.

TOCSY-HMQC : Spin systems are assigned by correlating NH resonances and the ^{15}N resonances of each residue with $^{13}C_{\alpha}$ and $C^{\alpha}H$ resonances of a sample dissolved in 90%H_2O/10%D_2O. The NH, ^{15}N, $C^{\alpha}H$ resonances are correlated using the double-resonance 3D ^{1}H-^{15}N TOCSY-HMQC experiment.

HCACO : This experiment links the carbonyl resonances with the intraresidue C^{α} and H^{α} resonances.[44]

HCCH-COSY and HCCH-TOCSY

The side chain spin systems are identified by correlating aliphatic ^{1}H resonances with their attached ^{13}C resonances using HCCH-COSY and HCCH-TOCSY experiments on samples dissolved in D_2O. The cross peaks in the spectra obtained using these experiments are not symmetric about the diagonal.

7. References

[1] Pharmacia Biotech catalogue

[2] Frangioni, J. V., Neel, B. G., (1993) Solubilization and purification of enzymatically active glutathione S-transferase (pGEX) fusion proteins, *Analytical Biochem.*, **210**, 179-187

[3] Bellacosa, A., Testa, J. R., Stall, S. P., Tsichlis, P. N., (1991) A retroviral oncogene, *akt*, encoding a serine-threonine kinase containing an SH2-like region, *Science*, **254**, 274-277

[4] Coffer, P. J., Woodgett, J. R., (1991) Molecular cloning and characterisation of a novel putative protein-serine kinase related to the camp-dependent and protein-kinase-C families, *Eur. J. Biochem*, **201**, 475-481

[5] Lodish H., Baltimore D., Berk A., Zipursky S. L., Matsudaira, P., Darnell J., (1995) Molecular Cell Biology, 3rd edition, Scientific American Books of New York

[6] Chen, D., Waters, S. B., Holt, K. H., Pessin J. E., (1996) SOS phosphorylation and disassociation of the Grb2-SOS complex by the ERK and JNK signalling pathways, *J. Biol. Chem.*, **271**, 6328-6332

[7] Pawson, T., Scott, J., (1997) Signalling through scaffold, anchoring and adaptor proteins, *Science*, **278**, 2075-2080

[8] Su, J. Y., Hodges, R. S., Kay, C. M., (1994) Effect of chain length on the formation and stability of synthetic alpha-helical-coiled coils, *Biochemistry*, **33**, 15501-15510

[9] Houston Jnr, M. E., Wallace, A., Bianchi, E., Pessi, A., Hodges, R. S., (1996) Use of a conformationally restricted secondary structural element to display peptide libraries, *J. Mol. Biol.*, **262**, 270-282

[10] Junius, F. K., Mackay, J. P., Bubb, W. A., Jenson, S. A., Weiss, A. S., King, G. F., (1995) Nuclear magnetic resonance characterisation of the Jun leucine zipper domain: unusual properties of coiled-coil interfacial polar residues, *Biochemistry*, **34**, 3164-3174

[11] Iivanainen, A., Coiled-coil motifs in proteins, http://cc.oulu.fin/~aiivanai/cc.html

[12] Tripet, B., Vale, R. D., Hodges, R. S., (1997) Demonstration of coiled-coil interactions within the kinesin neck region using synthetic peptides, *J. Biol. Chem.*, **272**, 8946-8956

[13] Hitchcock-DeGregori, S. E., An, Y., (1996) Integral repeats and a continuous coiled coil are required for binding of striated muscle tropomyosin to the regulated actin filament, *J. Biol. Chem.*, **271**, 3600-3603

[14] Hayward, C. P. M., Hassell, J. A., Denomme, G. A., Rachubinski, R. A., Brown, C., Kelton, J. G., (1995) The cDNA sequence of human endothelial cell multimerin, *J Biol. Chem.*, **270**, 18246-18251

[15] Stolton, C., Unpublished results.

[16] Lupas, A., Van Dyke, M., Stock, J., (1991) Predicting coiled coils from protein sequences, *Science*, **252**, 1162-1164

[17] Rost, B., Sander, C., (1993) Prediction of protein structure at better than 70% accuracy, *J. Mol. Biol.*, **232**, 584-599

[18] Jelesarov, I., Bosshard, H. R., (1996) Thermodynamic characterisation of the coupled folding and association of heterodimeric coiled coils (leucine zippers), *J. Mol. Biol.*, **263**, 344-358

[19] Moitra, J., Szilák, L., Krylov, D., Vinson, C., (1997) Leucine is the most stabilising aliphatic amino acid in the d position of a dimeric leucine zipper coiled coil, *Biochemistry*, **36**, 12567-12573

[20] DNASTAR, Inc 1994

[21] Stephens, L., Hawkins, P. T., Eguinoa, A., Cooke, F., (1996) A heterotrimeric GTPase-regulated isoform of PI3K and the regulation of its potential effectors, *Phil. Trans. R. Soc. Lond.* B, **351**, 211-215

[22] Buchman, V. L., personal communication

[23] Datta, K., Franke, T. F., Chan, T. O., Makris, A., Yang, S-I., Kaplan, D. R., Morrison, D. K., Golemis, E. A., Tsichlis, P. N., (1995) AH/PH domain-mediated interaction between Akt molecules and its potential role in Akt regulation, *Mol. Cell. Biol.*, **15**, 2304-2310

[24] Burgering, B. M. T., Coffer, P. J., (1995) Protein kinase B (c-Akt) in phosphatidylinositol-3-OH kinase signal transduction, *Nature*, 1995, **376**, 599-602

[25] Yao, R., Cooper, G. M., (1995) Requirement for phosphatidylinositol-3 kinase in the prevention of apoptosis by nerve growth factor, *Science*, **267**, 2003-2006

[26] Franke, T. F., Yang, S-I., Chan, T. O., Datta, K., Kazlauskas, A., Morrison, D. K., Kaplan, D. R., Tsichlis, P. N., (1995) The protein kinase encoded by the Akt proto-oncogene is a target of the PDGF-activated phosphatidylinositol 3-kinase, *Cell*, **81**, 727-736

[27] Marshall, C. J., (1995) Specificity of receptor tyrosine kinase signalling - transient versus sustained extracellular signal-regulated kinase activation, *Cell*, **80**, 179-185

[28] Hemmings, B. A., (1997) Akt signalling: linking membrane events to life and death decisions, *Science*, **275**, 628-630

[29] Kaplan, D. R., Miller, F. D., (1997) Signal transduction by the neurotrophin receptors, *Curr. Opin. Cell Biol.*, **9**, 213-221

[30] Marte, B. M., Downward, J., (1997) PKB/Akt: Connecting phosphoinositide 3-kinase to cell survival and beyond, *Trends Biochem. Sci.*, **22**, 355-358

[31] Bos, J. L., (1995) A target for phosphoinositide 3-kinase: Akt/PKB, *TIBS*, **20**, 441-442

[32] Cross, D. E. A., Alessi, D. R., Cohen, P., Andjelkovich, M., Hemmings, B. A., (1995) Inhibition of glycogen synthase kinase-3 by insulin mediated by protein kinase B, *Nature*, **378**, 785-789

[33] Dudek, H., Datta, S. R., Franke, T. F., Birnbaum, M. J., Yao, R., Cooper, G. M., Segal, R. A., Kaplan, D. R., Greenburg, M. E., (1997) Regulation of neuronal survival by the serine-threonine protein kinase Akt, *Science*, **275**, 661-665

[34] Franke, T. F., Kaplan, D. R., Cantley, L. C., Toker, A., (1997) Direct regulation of the Akt proto-oncogene product by phosphatidylinositol-3, 4-biphosphate, *Science*, **275**, 665-668

[35] Rodriguez-Viciana, P., Marte, B. M., Warne, P. H., Downward, J., (1996) Phosphatidylinositol 3 kinase: one of the effectors of Ras, *Phil. Trans. R. Soc. Lond.* B, **351**, 225-232

[36] Protein Purification Methods,(1989) Ed. Harris, E. L. V., Angal, S., IRS Press

[37] Kay, L. E., Ikura, M., Tschudin, R., Bax, A., (1990) Three-dimensional triple-resonance NMR spectroscopy of isotopically enriched proteins, *J. Mag. Resn.*, **89**, 496-514

[38] Zhao, D., Jardetzky, O., (1994) An assessment of the precision and accuracy of protein structures determined by NMR, *J. Mol. Biol.*, **239**, 601-607

[39] Bax, A., Ikura, M., Kay, L. E., Torchia, D. A., Tschudin, R., (1990) Comparison of different modes of two-dimensional reverse-correlation NMR for the study of proteins, *J. Mag. Resn*, **86**, 304-318

[40] Sattler, M., Fesik, S. W., (1996) Use of deuterium labelling in NMR: overcoming a sizeable problem, *Structure*, **4**, 1245-1249

[41] Yamazaki, T., Lee, W. L., Arrowsmith, C. H., Muhandiram, D. R., Kay, L. E., (1994) A suite of triple resonance NMR experiments for the backbone assignment of ^{15}N, ^{13}C, ^{2}H labelled proteins with high sensitivity, *J. Am. Chem. Soc.*, **116**, 11655-11666

[42] Venters, R. A., Huang, C. C., Farmer II, B. T., Trolard, R., Spicer, L. D., Fierke, C. A., (1995) High level $^{2}H/^{13}C/^{15}N$ labelling of proteins for NMR studies, *J. Biomol. NMR*, **5**, 339-344

[43] Fulton, D. B., Ni, F., (1997) ROESY with water flip back for high-field NMR of biomolecules, *J. Mag. Resn*, **129**, 93-97

[44] Evans, J. N. S., (1995) Biomolecular NMR Spectroscopy, Oxford University Press